(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
21 March 2002 (21.03.2002)

PCT

(10) International Publication Number
WO 02/22819 A2

(51) International Patent Classification[7]: C12N 15/12, 15/85, C07K 14/46, A61K 31/713, 38/17, 48/00, A01K 67/027

(21) International Application Number: PCT/IB01/01662

(22) International Filing Date:
13 September 2001 (13.09.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
00810830.0 14 September 2000 (14.09.2000) EP

(71) Applicant *(for all designated States except US)*: **UNIVERSITÄT ZÜRICH** [CH/CH]; Rämistrasse 71, CH-8006 Zürich (CH).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **SONDEREGGER, Peter** [CH/CH]; Krattenturmstrasse 1, CH-8006 Zürich (CH). **HINTSCH, Gustav** [CH/CH]; Köschenrütistrasse 11, CH-8052 Zürich (CH). **KINTER, Jochen** [DE/CH]; Müllerstrasse 56, CH-8004 Zürich (CH). **MESKENAITE, Virginija** [LT/CH]; Huttenstrasse 30, CH-8006 Zürich (CH). **SCHRIMPF, Sabine** [DE/CH]; Rebbergstrasse 8, CH-8037 Zürich (CH). **VOGT, Lorenz** [CH/CH]; Langfurrenstrasse 26, CH-8623 Wetzikon (CH). **ZURLINDEN, Andreas** [CH/CH]; Probusweg 9, CH-8097 Zürich (CH).

(74) Agent: **E. BLUM & CO.**; Vorderberg 11, CH-8044 Zürich (CH).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*
— *with sequence listing part of description published separately in electronic form and available upon request from the International Bureau*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/22819 A2

(54) Title: CALCIUM BINDING PROTEINS

(57) Abstract: There is described a novel class of calcium binding proteins of the nervous system, in particular calsyntenin-1-3. Calsyntenin proteins are valuable agents in the treatment of disorders of the nervous system, in particular the central nervous system. They are very useful for the development of drugs for the treatment of disorders of the nervous system.

Calcium binding Proteins

Cross References to Related Applications

This application claims the priority of European patent application 00810830.0, filed September 14, 2000, the disclosure of which is incorporated herein by reference in its entirety.

TECHNICAL FIELD

The present invention concerns a novel class of calcium binding proteins predominantly expressed in the nervous system.

BACKGROUND ART

Nervous system related disorders, in particular central nervous system related disorders, are getting greater importance, be it due to the enhanced average age of the people, be it due to enhanced numbers of injured people due to the enhanced occurrence of potential dangers, be it due to enhanced occurrence of stress-related and other psychic disorders.

There is a great interest in obtaining more knowledge about the nervous system and disorders involving the nervous system, such as psychic disorders, such as pain development, regeneration of injured nerves etc., and in particular about healing such disorders or injuries, or at least ameliorating the state of a patient suffering from such disorders or injuries, and there is a great need for pharmaceutical and diagnostic preparation in said field. One approach to learn more about nervous system related disorders is the identification and characterisation of proteins involved in biochemical pathways of the nervous system. Calcium plays an important role in signalling pathways of the nervous system. Although a lot

of DNA sequences of nervous system active proteins have been published, still a lot of such proteins have not yet been detected. Furthermore, also for most of the known sequences, their activity is still unknown.

For the above mentioned reasons there is a great interest to identify and characterise proteins expressed in the nervous system and playing a role in Calcium signalling or storage.

DISCLOSURE OF THE INVENTION

In nervous system derived cDNA libraries, recently the DNA sequence encoding a protein with so far unknown activity has been published. It has now been found in the scope of the present invention that said protein plays an important role in the calcium signalling pathway and is only one member of a whole class of compounds with similar activity. Said nervous system active protein that in the scope of the present invention has been denominated calsyntenin-1, has been found to comprise a single-pass transmembrane segment with a large extracellular segment and a small (approximately 100 amino acids) cytoplasmic segment highly enriched in acidic amino acid residues.

Nothing has been known prior to the present invention about the cellular pattern of calsyntenin gene expression, the cellular and subcellular localization and the functional role of the calsyntenin proteins, and calsyntenin-related disorders.

Hence it is an object of the present invention to provide an isolated polypeptide for the use as a pharmaceutical comprising an amino acid sequence at least 50% identical to a sequence selected from the group consisting of:

a) a full length amino acid sequence selected from the group consisting of Seq. Id. No. 2 (Calsyntenin-

1), Seq. Id. No. 4 (Calsyntenin-2) and Seq. Id. No. 6 (Calsyntenin-3),

b) a polypeptide comprising at least one, preferably two, most preferably three of amino acid sequences selected from the group consisting of: sequence of residues from about 46 to about 165, sequence of residues 166 to about 257, sequence of residues from about 774 to about 861 and sequence of residues from about 881 to about 981 of Seq. Id. No. 2,

c) a polypeptide comprising at least one, preferably two, most preferably three of amino acid sequences selected from the group consisting of: sequence of residues from about 66 to about 158, sequence of residues from about 182 to about 259, sequence of residues from about 751 to about 834 and sequence of residues from about 854 to about 955 of Seq. Id. No. 4,

d) a polypeptide comprising at least one, preferably two, most preferably three of amino acid sequences selected from the group consisting of: sequence of residues from about 51 to about 142, sequence of residues from about 167 to about 244, sequence of residues from about 759 to about 845 and sequence of residues from about 869 to about 956 of Seq Id. No. 6,

e) a polypeptide comprising at least one, preferably two, most preferably three of amino acid sequences selected from the group consisting of: sequence of residues from about 881 to about 981 of Seq. Id. No. 2, sequence of residues from about 854 to about 955 of Seq. Id. No. 4 and sequence of residues from about 869 to about 956 of Seq Id. No. 6,

and having calcium binding activity and/or having the capacity to bind to the Arp2/3 complex.

Polypeptides of the above defined group which comprise a ligand binding function are of particular interest, especially polypeptides comprising amino acid residues 46 to 165, residues from about 166 to about 257 of Seq.Id. No. 2; especially polypeptides comprising

amino acid residues from about 66 to about 158, residues from about 182 to about 259 of Seq. Id. No. 4; especially polypeptides comprising amino acid residues from about 51 to 142, residues from about 167 to about 244 of Seq. Id. No.6.

Polypeptides as defined above comprising a proteolytic cleavage site are preferred, especially a polypeptide comprising amino acid residues from about 774 to about 861 of Seq. Id. No. 2

especially a polypeptide comprising amino acid residues form about 751 to about 834 of Seq.Id. No. 4,

especially a polypeptide comprising amino acid residues from about 759 to 845 of Seq. Id. No. 6.

Polypeptides as defined above comprising a calcium binding domain of Calsyntenin-1 and/or Calsyntenin-2 and/or Calsyntenin-3 are preferred i.e. polypeptides comprising amino acid residues from about 881 to about 981 of Seq. Id. No. 2 and/or polypeptides comprising amino acid residues from about 854 to about 955 of Seq. Id. No. 4 and/or polypeptides comprising amino acid residues from about 869 to about 956 of Seq. Id. No. 6.

Preferred are polypeptide sequences that are at least 60% identical and more preferably more then 70% identical to an amino acid sequence selected from the above defined group.

Another object of the present invention is an isolated polypeptide for the use as a pharmaceutical comprising an amino acid sequence selected from sequences comprising a stretch of at least 100 amino acids with a minimal identity percentage of 50%, preferably 55% and more preferably 60% to an amino acid sequence selected from the group consisting of Seq. Id. No. 2, Seq. Id. No. 4 and Seq. Id. No. 6 and said sequences having calcium binding activity and/or the capacity to bind to the Arp2/3 complex.

The polypeptide of the present invention is preferably a transmembrane protein which is expressed predominantly in cells of the nervous system, and which is more preferably expressed in neurons.

The polypeptide of the present invention is preferably localised in a postsynaptic membrane of synapses, more preferably localized in a membrane of a spine apparatus of spine synapses and/or in a membrane of subsynaptic endoplasmatic reticulum of shaft synapses. Of particular interest are proteins having their major calcium-binding domain in the cytoplasmic compartment. Preferred are polypeptides which are expressed in tumors and other preferred polypeptides have at least one binding site for the Arp2/3 complex. Said Arp2/3 binding site is a conserved acidic amino acid sequence motive comprising a conserved tryptophan and encompasses but is not limited to e.g. the sequence motives MDWDDS and LEWDDS (amino acid sequence given in single letter code).

The polypeptides of the present invention or fragments thereof which have a minimal length of about 50 amino acids are suitable for the use as a tool for the development of a pharmaceutical.

Another object of the present invention is an isolated nucleotide sequence or a partial sequence thereof encoding a polypeptide of the present invention for the use as pharmaceutical.

Another object of the present invention is an isolated nucleotide sequence encoding a polypeptide of the present invention or a fragment thereof which has, due to at least one point mutation, insertion or deletion lost its function.

A further object of the present invention is an isolated nucleotide sequence encoding a polypeptide of the present invention or a fragment thereof, respectively which has, due to at least one point mutation, insertion or deletion lost its function for the use as a diagnostic tool. Such sequences, usually have not more than 25 dif-

ferences to the active segment within the Calcium binding intracellular region and/or the extracellular segment comprising the protease recognition site.

A mutated region or a region flanking the mutated region of the nucleotide sequence encoding a polypeptide of the present invention is e.g. useful for the design of primers or nucleotide probes that can be used in a diagnostic test to detect mutated DNA isolated from e.g. human tissue. Such test is e.g. suitable to predict whether cells are likely to undergo transformation leading to cancer development. The terms primer or nucleotide probe as used herein include oligonucleotide sequences comprised of ten or more deoxyribonucleotides or ribonucleotides.

DNA sequences of the present invention shall be understood to also include splice variants and DNA sequences that hybridize under stringent conditions to the sequences selected from the group consisting of Seq. Id. No. 1 (Calsyntenin-1), Seq. Id. No. 3 (Calsyntenin-2) and Seq. Id. No. 5 (Calsyntenin-3). Under stringent conditions hybridizing sequences in general are sequences with at least about 80 % identity, preferably about 90 % identity and most preferred 98 % identity. Said sequences comprise sequences encoding amino acid sequences having calcium binding activity as well as such sequences that encode amino acid sequences without calcium binding activity, in particular such sequences that for small defects have lost said activity. Small defect usually means a point mutation, an insertion or a deletion in said sequences. Sequences of the present invention also comprise sequences encoding amino acid sequences spanning over the proteolytic cleavage site(s) in the extracellular segment of the coded protein as well as such sequences that encode amino acid sequences without proteolytic cleavage site(s), in particular such sequences that for small defects have lost said cleavage site(s). The sequences of the present invention also comprise sequences encoding

the amino acid sequences of the proteolytically released fragments, as well as such sequences that encode amino acid sequences with mutations in the proteolytically released fragment. The term allele as used herein is intended to include sequences that differ by one or more nucleotide substitutions, additions or deletions, usually at most 20 differences in activity providing regions.

Another object of the present invention is the use of the polypeptides or the nucleotide sequences of the present invention or fragments thereof as a tool for the development of pharmaceuticals and as a tool for the screening of pharmaceutical agents.

The present invention furthermore concerns pharmaceutical compositions, that comprise such a DNA sequence and/or a polypeptide or a fragment thereof as defined above.

A pharmaceutical composition of the present invention can also comprise as at least one active substance (ingredient) a protein as defined above.

A pharmaceutical composition can furthermore comprise at least one further active compound, e.g. a compound that increases or reduces the calcium binding activity of said above defined protein, or that increases or decreases the amount of such a protein at its place of action in the body, or that prolongs or shortens the time of presence of such a protein at its place of action in the body.

The present invention furthermore encompasses a pharmaceutical composition that comprises as an at least one active compound a substance which enhances or inhibits the transcription of a mRNA derived from a DNA as defined above, or in that it enhances or inhibits the translation of such a DNA.

The present invention concerns as well a pharmaceutical composition, that comprises as an at least one active compound a compound that reduces or increases the calcium binding activity of a protein as defined

above, or that increases or decreases the amount of such a protein at its place of action in the body, or that shortens or prolongs the time of presence of such a protein at its place of action in the body.

Another object of the present invention are proteins having a sequence as specified in Seq. Id. No. 4 and homologues of said sequence comprising proteins which have at least 60% identity with said sequences.

Another object of the present invention are proteins having a sequence as specified in Seq. Id. No. 6 and homologous of said sequence comprising proteins which have at least 98.5 % identity with said sequence.

Another object of the present invention are nucleotide sequences which encode a protein as specified in Seq. Id. No. 4 or Seq. Id. No. 6. The coding sequence of the nucleotide sequence comprises all sequences encoding the amino acid sequence of Seq. Id. No. 4 or Seq. Id. No. 6 or homologues thereof. Also included are partial sequences of the nucleotide sequences described above. For instance, the calsyntenin encoding sequence preferably has a sequence at least 70 % similarity to the nucleotide sequence encoding the amino acid sequence of Seq. Id. No. 4 or Seq. Id. No. 6.

The DNA sequences and/or proteins defined above are suitable for the use in screening assays and/or the treatment of disorders, preferably nervous system disorders, more preferably of the central nervous system, most preferably the brain e.g. due to lack of calcium binding activity, or due to excessive calcium binding activity or due to perturbed processing of intracellular calcium signals and in particular in order to prevent, ameliorate or cure disorders of the nervous system caused due to lack of cleavage or miscleavage or excessive cleavage of a protein of the present invention induced by at least one protease, in particular proteases selected from the group consisting of tissue-type plasminogen activator, abbreviated as tPA, urokinase-type plasminogen

activator, abbreviated as uPA, or plasmin, or neurotrypsin or nuroserpin. Said disorders due to perturbed processing of intracellular calcium signals are preferably caused by perturbed processing of extracellular signals that regulate the cellular motility processes by means of regulating the activity of the Arp2/3 complex. Thus, a method for treating such diseases by use of a protein or a nucleotide sequence of the present invention is also encompassed.

Most preferably the present invention concerns DNA sequences or proteins for the minimization of the tissue destruction in stroke.

By a preparation comprising such DNA sequences or proteins, the minimization of the tissue destruction in stroke including brain infarction and ischemia, intracerebral hemorrhage, and subarrachnoid hemorrhage, as for example by exerting a protecting effect on the cells of the so-called penumbra zone surrounding the necrotic tissue, can be obtained.

Other disorders where an effective substance or preparation of this invention can be used, be it as pharmaceutical, be it as diagnostic agent, include as a suitable selection

the treatment of tissue destruction in traumatic brain injury, as for example by exerting a protective effect on the cells of the so-called penumbra zone surrounding the necrotic tissue,

the prevention, amelioration or cure of negative effects caused by neurodegenerative diseases, or neuroinflammatory diseases, as for example multiple sclerosis,

the reduction or prevention of negative effects on brain tissue caused by epileptic seizures,

the rescue of endangered neurons, as for example neurons endangered by hypoxia and ischemia, excitotoxicity, neuroinflammatory diseases and processes, epileptic seizures, and cancerous neoformations,

the axonal regeneration and/or restoration of synaptic integrity and functions,

the prevention, amelioration, or cure of retinal disorders, as for example retinal degeneration and retinal neoangiogenesis,

the cell death of cells of the nervous system, in particular a cell death in connection with damages of the nervous tissue, for example infarct of the brain and ischemic stroke, or hemorrhage of the brain, or trauma of the brain, and/or a cell death in connection with damages of the nervous tissue, which occur due to lack of oxygen or glucose or due to intoxication, and/or a cell death in connection with epileptic seizures, and/or a cell death in connection with neurodegenerative diseases and inherited genetic deficiencies of the nervous system,

the regeneration of injured, damaged, underdeveloped, or maldeveloped brain tissue and/or nervous tissue,

the reorganization of the brain or nervous areas that have remained intact after brain and/or nerve injuries or after the destruction or damage of brain areas,

the prevention, amelioration, or cure of pathological pain syndromes,

the amelioration or cure of disorders in the field of disorders of the psychic wellness, or the psychosomatic state of health, as for example nervosity or „inner unrest", disorders in the field of the emotional functions, as for example states of anxiety,

the prevention, amelioration or cure of psychiatric disorders, in particular psychiatric disorders in the field of schizophrenia and schizophrenia-like disorders, including chronic schizophrenia, chronic schizoaffective disorders, unspecific disorders, including acute and chronic schizophrenia of various symptomatologies, as for example severe, non-remitting „Kraepelinic"

schizophrenia, or as for example the DSM-III-R-prototype of the schizophrenia-like disorders, including episodic schizophrenic disorders, including delusionic schizophrenia-like disorders, including schizophrenia-like personality disorders, as for example schizophrenia-like personality disorders with mild symptomatics, including schizotypic personality disorders, including the latent forms of schizophrenic or schizophrenia-like disorders, including non-organic psychotic disorders, and/or in the field of the endogenic depressions or in the field of manic or manic-depressive disorders,

the treatment of tumors such as prevention or reduction of the growth, the expansion, the infiltration and the metastasis of primary and metastatic tumors, inhibition of the formation of new blood vessels or neoangiogenesis, in particular the treatment of brain tumors or tumors of the retina. Said tumors preferably involve in their growth, expansion, infiltration, metastasis and promotion of blood vessels or neoangiogenesis an enhanced activity of the Arp2/3 complex. Said enhanced activity of the Arp2/3 complex is preferably mediated by an abnormal or excessive or reduced regulatory function of one of the proteins of the present invention.

Said tumors preferably involve in their growth, expansion, infiltration, metastasis and promotion of blood vessels or neoangiogenesis at least one protease functionally connected with a polypeptide of the present invention. Said protease is preferably a member of one of the following families:

- Serine Protease family such as tissue-type plasminogen activator (tPA), urokinase-type plasminogen activator (uPA), plasmin, thrombin, neurotrypsin, neuropsin, elastases, cathepsin G,
- Matrix Metalloproteinases family such as collagenases, gelatinases, stromelysins, matrylisins,
- Cystein Proteases family such as cathepsin B and cathepsin D.

The present invention also concerns the amelioration of the learning and memory functions in healthy persons, as well as in persons with reduced learning and memory functions.

In one additional aspect, the present invention concerns a method for the production of proteins as defined above, that is characterized in that suitable host procaryotic and eucaryotic cells, in particular mammalian cells, are transfected with a DNA sequence as defined above in a vector ensuring the expression of said DNA sequence, and in that said transfected cells are cultured under suitable conditions allowing expression of said protein.

In another object the present invention relates to a synthetic or chemical method for the production of polypeptides, peptides or nucleic acid sequences representing at least part of the sequences of the present invention and having the ability to mimic or to block, respectively, the biological activity or calsyntenin, in particular the calcium binding activity.

The DNA sequences and/or the proteins defined above can furthermore be used as means for the screening of drugs against calsyntenin protein involving disorders, but also active ingredients such as transcription enhancers or reducers and translation enhancers or reducers and activity enhancers or reducers.

Another object of the present invention is a protease or proteases cleaving the proteins of the present invention in their extracellular segment.

Furthermore the present invention relates to cell extracts comprising a protease which cleaves a polypeptide of the present invention. The protease can have endogenous origin or can be the product of a heterologous expression construct transformed or transfected into said cells.

Another object of the present invention is a method for the identification of a compound or an agent

which modulates the activity of said proteases. Said method comprises contacting cells producing an active protease with a test compound and measuring changes in protease activity. In a preferred embodiment said cells are mammalian cells and the protease is expressed from a heterologous gene construct.

Furthermore, the present invention also comprises the use of a sequence as defined above as a means to produce antigens or as antigen for the production of antibodies.

Such antibodies can e.g. be antibodies that inhibit or promote the calcium binding function or antibodies that inhibit or promote the proteolytic cleavage of a protein as defined above or antibodies that can be used for immunohistochemical studies or diagnostic assays.

The present invention also regards transgenic animals comprising an exogenous DNA sequence as defined above. Such animals are suitable for the study of diseases and the test of active substances as defined above

Such animals are in particular non human mammals, such as mice.

Still a further aspect of the present invention concerns the use of a DNA sequence as defined above for the inactivation or the mutation of the corresponding endogenous gene by means of gene targeting techniques.

Such gene targeting techniques are for example the elimination of the gene in the mouse through homologous recombination or the replacement of the gene by a mutated form thereof.

A DNA sequence as defined above can, within the scope of the present invention, also be used for the preparation of a diagnostic preparation for the diagnosis of disorders due to defects or alterations in the genomic sequence comprising a coding sequence similar to but not identical with one of the coding sequences defined above.

The nucleic acid sequences of the present invention are of great interest in gene therapeutical applications in humans and in animals, as for example as parts of gene therapy vectors, such as biological and synthetic vectors, or as parts of artificial chromosomes.

Brief Description of the Drawings

The invention will be better understood and objects other than those set forth above will become apparent when consideration is given to the following detailed description thereof. Such description makes reference to the annexed drawings, wherein:

Figure 1A shows dissociated neurons from the ventral halves of E6 chicken spinal cords seeded in the central compartment of a cell culture system,

Figure 1B shows neurites of neurons from the ventral halves of E6 chicken spinal cords extending into the side compartment after 6 days of cultivation,

Figure 1C shows a compartmental cell culture system, the cell culture surface is subdivided into three compartments by a Teflon divider,

Figure 1D shows a fluorography of a two-dimensional SDS-PAGE gel of ^{35}S methionine labelled proteins released into the medium of both the central and the side compartments,

Figure 1E shows a fluorography of a two-dimensional SDS-PAGE gel of ^{35}S methionine labelled proteins released into the medium of both the central and the side compartments,

Figure 2 shows an alignment of amino acid sequences deduced from the single ORF in the human (hs), the mouse (mm) and the chicken (gg) cDNA of calsyntenin-1,

Figure 3 shows a demonstration of the calcium binding capacity of the cytoplasmic moiety of calsyntenin-1,

Figure 4A shows an expression pattern of calsyntenin-1 mRNA in a sagital section of an E18 mouse,

Figure 4B shows a an expression pattern of calsyntenin-1 mRNA in a coronal section of an adult mouse brain,

Figure 4C shows a Northern blot analysis of calsyntenin-1 mRNA in adult human tissues,

Figure 4D shows a Western blot analysis of human and chicken calsyntenin-1 protein,

Figure 4E shows a schematic drawing of the calsyntenin-1 protein,

Figure 5A shows synaptic localisation of calsyntenin-1 by immunohistochemical staining in a section of the hippocampus of an adult rat,

Figure 5B shows colocalization of calysntenin-1 with the synaptic marker synapthophysin,

Figure 5C shows colocalization of calsyntenin-1 with the α2 subunit of the synaptic marker $GABA_A$ receptor,

Figure 5D shows colocalization of calsyntenin-1 with the GluR2 subunit of the AMPA receptor,

Figure 6A shows an ultrastructural localization of calsyntenin-1 in the postsynaptic membrane of spine synapses,

Figure 6B shows an ultrastructural localization of calsyntenin-1 in the postsynaptic membrane of spine synapses,

Figure 6C an ultrastructural localization of calsyntenin-1 in the postsynaptic membrane of spine synapses,

Figure 6D shows an ultrastructural localization of calsyntenin-1 in the postsynaptic membrane of spine synapses,

Figure 6E an ultrastructural localization of calsyntenin-1 in the postsynaptic membrane of spine synapses,

Figure 6F an ultrastructural localization of calsyntenin-1 in the postsynaptic membrane of spine synapses,

Figure 6G an ultrastructural localization of calsyntenin-1 in the postsynaptic membrane of spine synapses,

Figure 7 shows localisation of calsyntenin-1 in synaptosomes, but not in postsynaptic densities,

Figure 8A shows ultrastructural localization of the transmembrane fragment of proteolytically cleaved calsyntenin-1 over the spine apparatus of spine synapses,

Figure 8B shows ultrastructural localization of the transmembrane fragment of proteolytically cleaved calsyntenin-1 over the spine apparatus of spine synapses,

Figure 8C shows ultrastructural localization of the transmembrane fragment of proteolytically cleaved calsyntenin-1 over the spine apparatus of spine synapses,

Figure 8D shows ultrastructural localization of the transmembrane fragment of proteolytically cleaved calsyntenin-1 over the spine apparatus of spine synapses,

Figure 8E shows ultrastructural localization of the transmembrane fragment of proteolytically cleaved calsyntenin-1 over the spine apparatus of spine synapses,

Figure 9 shows a diagram of the protease dependent translocation of the postsynaptic Ca^{2+} binding of calsyntenin-1

Figure 10A shows the localization of calsyntenin-1 in growth cones of cultured hippocampal neurons by indirect immunofluorescence staining,

Figure 10B identifies one of the neuronal processes shown in Figure 10A as an axon by indirect immunoflurescence staining with an antibody against the axonal marker protein Tau 1,

Figure 10C shows the localization of calsyntenin-1 in growth cone of cultured hippocampal neurons by indirect immunofluorescence at higher mangification,

Figure 11 shows the production of the full-length form of calsyntenin 1 and the N-terminal secreted fragment of cleaved calsyntenin by Western Blotting with antibodies against calsyntenin-1 (R63 and R71).

Figure 12A shows the expression of calsyntenin-3 mRNA in different human organs by Northern blotting,

Figure 12B shows the expression of calsyntenin-3 mRNA at cellular resolution in a horizontal section through a brain of an adult mouse and

Figure 12C shows the expression of calsyntenin-3 mRNA at cellular resolution in a parasagittal section through a brain of an adult mouse.

Figure 13 demonstrates the binding of the Arp2/3 complex to the cytoplasmic part of calsyntenin-1. Bovine brain extract was passed over a column containing bound GST-Cst$_C$ fusion protein. The proteins collected in the flow-through fraction, in the wash fractions, and in the elution fractions were separated by SDS-PAGE and stained with silver staining (upper panel). The same protein fractions were also electrotransferred after SDS-PAGE to nitrocellulose paper and the presence of Arp2/3 complex was visualized by immunodetection using as a first antibody a commercially available antibody directed against the Arp3 subunit of the Arp2/3 complex (lower panel).

MODES FOR CARRYING OUT THE INVENTION

Calsyntenin proteins are known to be expressed predominantly in the brain; the gene expression in the brain takes place nearly exclusively in the neurons.

As representatives of the novel class of calcium binding proteins the isolation and characterisation of calsyntenin-1, 2 and 3 are further described.

The coded peptide of calsyntenin-1 has a length of 1009 amino acids and contains a signal peptide of 28 amino acids. The mature protein is composed of 981 amino acids. The extracellular segment comprises 860 amino acids, the transmembrane segments has 21 amino acids, and the cytoplasmic segment has 99 amino acids.

A particularly interesting function of calsyntenin-1 is found in the segment forming the calcium binding cytoplasmic region.

Calsyntenin-1 has a cytoplasmic segment that is highly enriched in acidic amino acid residues and has the capacity of binding calcium ions. The cytoplasmic segment of the calsyntenin-1 functions as high-capacity, low-affinity calcium binding structure and it also contains high-affinity binding sites for calcium..

By this function, calsyntenin-1 retains calcium in the subsynaptic zone of excitatory and inhibitory synapses of the central and the peripheral nervous system. By this feature, calsyntenin-1 mediates the accumulation of calcium in the zone beneath the postsynaptic membrane and thereby modulate the calcium-mediated synaptic functions. By these functions, calsyntenin-1 maintains elevated calcium concentrations in the zone beneath the postsynaptic membrane and thereby, prolong the calcium signals in the zone beneath the postsynaptic membrane. An interesting aspect of calsyntenin-1 is its removal from the postsynaptic membrane by an endocytic process that follows the proteolytic cleavage within the extracellular segment. Therefore, calsyntenin-1 is subject to dynamic regulations by proteolytic cleavage by at least one synaptic protease.

In the scope of the present invention it could now be shown that calsyntenin-1 has also an *in vivo* activity making it a very useful tool for the diagnostic and therapy of protease involving disorders, in particular of the the nervous system, more particular of the central nervous system.

It is known that the expression of calsyntenin-1 during neural development starts at the beginning of the time range in which restructuration processes of synapses are observed, that in the adult nervous system, their expression is predominant in brain regions in which synapse plasticity occurs, and that a particularly high expression of calsyntenin-1 is found in the cerebral cortex, the hippocampus, and the amygdala of the mouse.

In the deeper structures of the brain, in the brain stem, and in the spinal cord of the adult mouse, a weaker expression of the calsyntenin-1 is found.

In the adult peripheral nervous system, calsyntenin-1 is also expressed, for example in the sensory ganglia neurons.

The gene expression pattern of calsyntenin-1 in the brain is extremely interesting, because these molecules are expressed in the adult nervous system predominantly in neurons of those regions that are thought to play an important role in learning and in memory functions.

The gene expression pattern of calsyntenin-1 in the cerebral cortex is extremely interesting, because a reduction of the cellular differentiation in the cerebral cortex has been found to be associated with schizophrenia.

Another prominent characteristic of calsyntenin-1 consists therein that it is secreted by neurons.

This fact - together with the function as a calcium binding protein and the expression pattern in the developing and adult brain - suggests that the calsyntenin-1 plays a role in the regulation of the calcium-mediated signals in brain areas which are involved in the processing and storage of learned behaviors, learned emotions, or memory contents.

Together with the recently found evidence for a role of extracellular proteases, in particular tissue-type plasminogen activator, in neural plasticity (see

Frey et al., J. Neurosci. 16, pages 2057-2063, 1996; Huang et al., Proc. Natl. Acad. Sci. USA 93, pages 699-704, 1996), the expression pattern allows the assumption that the calcium binding activity of calsyntenin has a role in learning and memory operations, for example operations which are involved in the processing and storage of learned behaviors, learned emotions, or memory contents.

The fact that calsyntenin-1 is a substrate of proteases is particularly interesting, because for example the protease tissue-type plasminogen activator (tPA) has been found to play a role in the pathogenesis of neuronal cell damage or neuronal cell death in the context of excitotoxin-induced epileptic seizures (see Tsirka et al., Nature 377, pages 340-344, 1995).

The gene expression pattern of the calsyntenin-1 in the spinal cord and in the sensory ganglia is interesting, because these molecules are expressed in the adult nervous system in neurons of those brain regions that are thought to play a role in the processing of pain, as well as in the pathogenesis of pathological pain.

Calsyntenin-1 was found in connection with a study aimed at discovering proteins that are secreted from axons of neurons (see Stoeckli et al., Eur. J. Biochem. 180, pages 249-259, 1989). Their preparation has now been described in several papers that are herein comprised by reference (see Osterwalder et al., EMBO J. 15, pages 2944, 1996; Schrimpf et al. Human Neuroserpin (PI12): cDNA Cloning and Chromosomal Localization to 3q26, *Genomics, Vol. 39,* pages 1-8 (1997).

This procedure for the cloning can also be used for the isolation of homologous sequences of other species, such as mouse, rat, rabbit, guinea pig, cow, sheep, pig, primates, birds, zebra fish (Brachydanio rerio), Drosophila melanogaster, Caenorhabditis elegans

etc. Such sequences are preferred for the veterinary use in order to avoid incompatibility reactions.

The coding nucleotide sequences obtained e.g. by the above described methods can be used for the production of proteins with the coded amino acid sequences as defined above.

The coding sequences of calsyntenin genes can also be used as starting sequence for the isolation of alleles and splice variants, or parts thereof, can be used as probes for the isolation of the genes corresponding to said sequences. For example the polymerase chain reaction and the nucleic acid hybridization technique can be used for this purpose.

The coding sequences of the calsyntenin genes can be used as starting sequences for so-called "site-directed mutagenesis", in order to generate nucleotide sequences encoding proteins as defined above, in particular those shown in Seq. Id. No. 1, 3 and 5, or parts thereof, but whose nucleotide sequence is degenerated with respect to the sequences shown in Seq. Id. No. 1, 3 and 5 due to use of alternative codons. Such mutagenesis can be desired dependent of the host cells used for the expression of the protein of interest.

The coding sequences disclosed in this invention, can be used as starting sequences for the production of sequence variants exhibiting altered function by means of so-called site-directed mutagenesis. Such altered functions can e.g. provide for proteins with longer lifetime, i.e. slower degradation, enhanced activity etc.

The coding sequences can be used for the production of vectors for use in gene therapy and cell engineering.

The coding sequences can be used for the generation of transgenic animals overexpressing the coding and the coded sequences of the present invention.

The coding sequences can be used for the diagnostics of disorders in the gene corresponding to the sequences of the present invention.

The amino acid sequences coded by the above described nucleic acid sequences can be used as active substances, as antigens for the production of antibodies, and as targets for drug development.

In a further aspect, the present invention relates to the use of the polypeptides or the nucleotide sequences of the present invention or fragments thereof as a tool for the development of pharmaceuticals and as a tool for the screening of pharmaceutical agents, in particular screening assays for compounds binding a protein of the present invention. A preferred target sequence of the proteins of the present invention for the binding of such molecules is the extracellular part of the proteins of the present invention, in particular the domain/site showing blue sepharose binding capacity. Another target sequence for the binding of such molecules is the intracellular Arp2/3 complex binding domain, in particular a sequence comprising the motives MDWDDS.. and..LEWDDS (amino acid sequence given in single letter code), of the proteins of the present invention.

Suitable in vitro assays for the identification of compounds which have an effect on the activity and/or stability and/or expression of the proteins of the present invention are for example in vitro assays employing biochemical or biophysical tests able to detect specific protein/ligand interactions and include e.g. MS/NMR as described in Moy et al. Anal. Chem. 73(3):571-81, 2001, high-throughput nuclear magnetic resonance-based screeening as described in Hajduk PJ. J. Med. Chem. 42(13): 2315-7, 1999 or mass spectrometry-based strategies as described in Kaur S., J. Protein Chem., 16(5): 505-11, 1997 which are incorporated herein by reference in its entirety.

Said aspect of the present invention is based on the findings that calsyntenin-1, and most likely also the family members calsyntenin-2 and calsyntenin-3, are capable of binding the Arp2/3 complex. Binding of the cytosolic segment of calsyntenin-1 to the Arp2/3 complex indicates a role of the calsyntenin protein family in the regulation of cell motility. While studying the scientific literature dealing with interactions between cell surface proteins and the cellular cytoskeleton, we found that the cytoplasmic part of all proteins of the calsyntenin family (i.e. calsyntenin-1, calsyntenin-2, and calsyntenin-3) contains at least one intriguing motif of conserved amino acid sequences containing a conserved tryptophan. This motif is highly similar to conserved acidic amino acid motifs with a conserved tryptophan found in the Arp2/3-binding domain of most, if not all, of the currently known activators of the Arp2/3 complex. In the cytoplasmic segment of calsyntenin-1, this motif is found twice, once with the amino acid sequence ..MDWDDS.. and once with ..LEWDDS.. (amino acid sequence given in single letter code). The cytoplasmic sequence of calsyntenin-2 contains one ..MDWDDS.. and one ..LEWDDS... The cytoplasmic segment of calsyntenin-3 contains a single motif of this kind, namely ..LFWDDS... The Arp2/3 complex plays a central role in the regulation of actin-based cellular motility, by regulating actin filament growth and branching (for reviews see: Borisy and Svitkina, Curr. Opin. Cell Biol. 12: 104-112, 2000; Pantaloni et al., Science 292: 1502-1506; Higgs and Pollard, Annu. Rev. Biochem., 70: 649-676, 2001; and references therein). Arp2/3 activators containing a similar motif with conserved acidic amino acids and a tryptophan include human WASP (Abbreviation for: Wiscott Aldrich Syndrome Protein), the related human N-WASP, the human Scar/WAVE1 proteins, and cortactin, exhibiting the sequences ..DDEWDD, ..DDEWED and ..EVDWLE, and ..ADDWET.., respectively (for WASP, N-WASP, and Scar/WAVE1 see Higgs and

Pollard, Annu. Rev. Biochem. 70: 649-676, 2001; for cortactin see Uruno et al., Nature Cell Biol. 3: 259-266, 2001). The importance of the conserved tryptophan and the adjacent acidic amino acids for Arp2/3 binding and Arp2/3 function in actin polymerization has been demonstrated by site-directed mutagenesis of cortactin (Uruno et al., Nature Cell Biol. 3: 259-266, 2001). Site directed mutagenesis of both the tryptophan and the two amino acid residues preceeding the tryptophan in the sequence ..ADDWET.. of cortactin resulted in the loss of Arp2/3 binding and Arp2/3-mediated actin polymerization. All these Arp2/3 activator proteins residene in the cytoplasm. They link intracellular signals derived from the interaction of transmembrane receptors with their extracellular regulators, such as growth factor, cytokines, etc., to activation of the Arp2/3 complex. A crucial intermediate step in the signaling cascade from activated transmembrane receptors to the activation of the Arp2/3 activators has been attributed to the small GTP-binding proteins of the Rho family (for a review: Takai et al., Physiol. Rev. 81:153-207, 2001). Activated Arp2/3 complex initiates the generation of new actin filaments and the branching of pre-existing actin filaments (for reviews see: Borisy and Svitkina, Curr. Opin. Cell Biol. 12: 104-112, 2000; Pantaloni et al., Science 292: 1502-1506; Higgs and Pollard, Annu. Rev. Biochem., 70: 649-676, 2001; and references therein). As a result of the enhanced cytoskeletal dynamics, the cells generate and/or retract plasma membrane protrusions, such as filopodia and lamellipodia (Borisy and Svitkina, Curr. Opin. Cell Biol. 12: 104-112, 2000). This enhanced activity translates into an enhanced exploratory activity of the growth cones, the growing tip of the axons extending from neurons as well as enhanced axon growth and pathfinding activity (Hu and Reichardt, Neuron 22, 419-422, 1999; Suter and Forscher, Curr. Opin. Neurobiol. 8: 106-116, 1998; Dickson, Curr. Opin. Neurobiol. 11: 103-110, 2001). In the dendritic spines of neurons of

the central nervous system, the enhanced dynamics of actin filaments results in an increase in motility, which in turn may regulate the morphological shape and the electrical properties of the spine. As a consequence, the postsynaptic response to presynaptic signals may be altered (Segal et al., Trends Neurosci. 23: 53-57, 2000; Halpain, Trends Neurosci. 23: 141-146, 2000; Matus, Science 290: 754-758, 2000; Scott and Luo, Nature Neurosci. 4: 359-365, 2001). In non-neuronal cells, the intensification of actin filament dynamics induced via Arp2/3 activation results in an enhanced cell motility that is accompanied by enhanced formation of membrane protrusions, such as lamellipodia, and enhanced migratory activity (Holt and Koffer, Trends Cell Biol. 11: 38-47, 2001; Mullins, Curr. Opin. Cell Biol. 12: 91-96, 2000; Prokopenko et al., J. Cell Biol. 148: 843-848, 2000). A dysregulated signalling from the cell surface to the cytoskeleton changes the migratory activity of tumor cells that is linked to their enhanced capacity for invasive growth and metastasis (Radisky et al., Seminars Cancer Biol. 11:87-95, 2001; Kassis et al., Seminars Cancer Biol. 11:105-119, 2001; Condeelis et al., Seminars Cancer Biol. 11:119-128, 2001; Price and Collard, Seminars Cancer Biol. 11:167-173, 2001).

The present invention provides methods to evaluate the activity of a compound to selectively regulate synaptic calcium signals. The rationale of the screening approach presented here is based on the immunoelectron microscopic studies presented herein. In these studies we found that full-length calsyntenin-1 is almost exclusively located in and beneath the postsynaptic membrane, whereas the transmembrane fragments generated by proteolytic cleavage is translocated to the membranes of the so-called spine apparatus. The complete absence of full-length calsyntenin-1 from the spine apparatus indicates that only proteolytically cleaved calsyntenin-1 is internalized. Obviously, the proteolytic cleavage is a

prerequisite for the internalization of the transmembrane segment of calsyntenin-1. As a result of the proteolytic cleavage of calsyntenin-1 and the subsequent internalization of its transmembrane fragment, the amount of calsyntenin-1 in the postsynaptic membranes is decreased. As a consequence, the regulatory influence of the cytoplasmic segment of calsyntenin-1 on synaptic calcium signaling is decreased.

Based on these characteristics, the proteolytic cleavage of calsyntenin-1 in its extracellular segment correlates with a reduction of the calsyntenin-1-mediated calcium-binding capacity beneath the postsynaptic membrane. Therefore, the extent of the proteolytic cleavage of calsyntenin-1 provides a correlate for the calsyntenin-1-mediated regulation of postsynaptic calcium signals. This link between proteolytic cleavage of calsyntenin-1 and the calsyntenin-1-mediated regulation of synaptic calcium signals can be exploited for the establishment of a relatively simple assay for testing a compound for its potential activity as a modulator of synaptic calcium signalling. This assay comprises contacting a calsyntenin protein expressing and synapse-forming neuronal cell culture or a synaptosomal or synaptoneurosomal preparation with a preselected amount of the compound in a suitable culture medium or buffer. After a suitable period of incubation, the progress of the proteolytic cleavage reaction of a full-length calsyntenin protein is assessed by measuring the decrease in the full-length form of calsyntenin and the increase in the two cleavage products. Measuring the degradation of a full-length calsyntenin protein and/or the generation of cleavage products of a calsyntenin protein by said neuronal cell culture or said synaptosomes or synaptoneurosomes, as compared to a control, will provide a measure for the efficiency of a compound in modulating endocytosis of a calsyntenin protein and, thus, the translocation of the calsyntenin--binding domain from the zone beneath the post-

synaptic membrane of the spine apparatus and, thus, the modulation of postsynaptic calcium signals.

More specifically, the present invention provides a method of determining the ability of a compound to influence the cleavage of a calsyntenin protein in the extracellular moiety. A typical experiment consists in:

a) preparation of a synapse-forming neuronal cell culture (e.g. dissociated hippocampal culture from mouse or rat brain: Goslin et al., 1998, Cultering nerve cells, 2nd Ed., MIT Press Cambridge, MA; or, alternatively, preparation of synaptosomes or synaptoneurosomes from brains of rodents (mouse or rat) or birds (chicken or pigeon) using established protocols (for synaptosomes: Phelan and Gordon-Weeks, 1997, Neurochemistry, A practical approach, 2nd Ed.; for synaptoneurosomes: Hollingsworth et al., 1985, J. Neurosci. 5, 2240-2253)

b) addition of the compound to the culture medium of the synapse-forming neuronal cell culture or to the buffer containing the suspended synaptosomes or synaptoneurosomes.

c) separation of the cellular or synaptosomal or synaptoneurosomal proteins by sodiumdodecylsulfate polyacrylamide gel electrophoresis.

d) visualization of a full-length calsyntenin and fragments thereof by Western blot analysis.

e) measurement of the relative amounts of full-length and cleaved calsyntenin.

f) comparison of the relative amounts of cleaved and uncleaved calsyntenin with the relative amounts of cleaved and uncleaved calsyntenin obtained under the control condition.

The present invention provides simple in vitro systems for the screening of drug actions on synaptic calcium signalling, which will be useful for the development of drugs that selectively modulate synaptic calcium signal without producing side effects due to modulation of nonsynaptic calcium signals. Assays can be performed

on living synapse-forming cultures of mammalian or avian neurons or on isolated mammalian or avian synapses (so-called synaptosomes or synaptoneurosomes), which can be cultivated or prepared, respectively, with relative ease. The assessment of the proteolytic cleavage of a calsyntenin by Western blot analysis is a relatively simple procedure as well. Thus, the assay is suited for high-throughput screening of a large number of compounds.

The invention also relates to methods for the identification of genes, termed "pathway genes", which are associated with a calsyntenin gene product or with the biochemical pathways which extend therefrom. "Pathway gene", as used herein, refers to a gene whose gene product exhibits the ability to interact with a calsyntenin gene product.

Any method suitable for detecting protein-protein interactions may be employed for identifying pathway gene products by identifying interactions between gene products and a calsyntenin gene product. Such known gene products may be an intracellular, a transmembranal, or an extracellular protein. Those gene products which interact with such known gene products represent pathway gene products and the genes which encode them represent pathway genes.

Among the traditional methods which may be employed are co-immunoprecipitation, crosslinking and co-purification through gradients or chromatographic columns. Utilizing procedures such as these allows for the identification of pathway gene products. Once identified, a pathway gene product may be used, in conjunction with standard techniques, to identify its corresponding pathway gene. For example, at least a portion of the amino acid sequence of the pathway gene product may be ascertained using techniques well known to those of skill in the art, such as via the Edman degradation technique (see, e.g., Creighton, 1983, Proteins: Structures and Molecular Principles, W. H. Freeman & Co., New York, pp.34-

49). The amino acid sequence obtained may be used as a guide for the generation of oligonucleotide mixtures that can be used to screen for pathway gene sequences. Screening made be accomplished, for example by standard hybridization or PCR techniques. Techniques for the generation of oligonucleotide mixtures and screening are well-known. (See, e.g., Ausubel et al., eds., 1987-2000, Current Protocols in Molecular Biology, John Wiley & Sons, Inc. New York, and PCR Protocols: A Guide to Methods and Applications, 1990, Innis, M. et al., eds. Academic Press, Inc., New York).

Additionally, methods may be employed which result in the simultaneous identification of pathway genes which encode the protein interacting with a calsyntenin gene product. These methods include, for example, probing expression libraries with a labeled calsyntenin protein, using this protein in a manner similar to the well known technique of antibody probing of lambda gt11 libraries.

One such method which detects protein interactions in vivo, the two-hybrid system, is described in detail for illustration only and not by way of limitation. One version of this system has been described (Chien et al., 1991, Proc. Natl. Acad. Sci. USA, 88:9578-9582) and is commercially available from Clontech (Palo Alto, Calif.).

Briefly, utilizing such a system, plasmids are constructed that encode two hybrid proteins: one consists of the DNA-binding domain of a transcription activator protein fused to a known protein, and the other consists of the activator protein's activation domain fused to an unknown protein that is encoded by a cDNA which has been recombined into this plasmid as part of a cDNA library. The plasmids are transformed into a strain of the yeast Saccharomyces cerevisiae that contains a reporter gene (e.g., lacZ) whose regulatory region contains the activator's binding sites. Either hybrid protein

alone cannot activate transcription of the reporter gene: the DNA-binding domain hybrid because it does not provide activation function and the activation domain hybrid because it cannot localize to the activator's binding sites. Interaction of the two proteins reconstitutes the functional activator protein and results in expression of the reporter gene, which is detected by an assay for the reporter gene product.

The two-hybrid system or related methodology may be used to screen activation domain libraries for proteins that interact with a calsyntenin gene product, herein also called the known "bait" gene protein. Total genomic or cDNA sequences may be fused to the DNA encoding an activation domain. Such a library and a plasmid encoding a hybrid of the bait gene protein fused to the DNA-binding domain may be cotransformed into a yeast reporter strain, and the resulting transformants may be screened for those that express the reporter gene. These colonies may be purified and the library plasmids responsible for reporter gene expression may be isolated. DNA sequencing may then be used to identify the proteins encoded by the library plasmids.

For example, and not by way of limitation, the bait gene may be cloned into a vector such that it is translationally fused to the DNA encoding the DNA-binding domain of the GAL4 protein.

A cDNA library of the cell line from which proteins that interact with bait gene are to be detected can be made using methods routinely practiced in the art. According to the particular system described herein, for example, the cDNA fragments may be inserted into a vector such that they are translationally fused to the activation domain of GAL4. This library may be co-transformed along with the bait gene-GAL4 fusion plasmid into a yeast strain which contains a lacZ gene driven by a promoter which contains the GAL4 activation sequence. A cDNA encoded protein, fused to the GAL4 activation domain, that

interacts with bait gene will reconstitute an active GAL4 protein and thereby drive expression of the lacZ gene. Colonies which express lacZ may be detected by their blue color in the presence of X-gal. The cDNA may then be purified from these strains, and used to produce and isolate the bait gene-interacting protein using techniques routinely practiced in the art.

Another method for discovering pathway genes is eucaryotic expression cloning. The expression-cloning method allows the isolation of a cDNA encoding a molecule that physically interacts with the protein of interest from a cDNA library contained in a eucaryotic expression vector. Over the past years it has emerged as a powerful method to identify the binding partners of many different proteins and other molecules (see Simmons, 1993, Cloning cell surface molecules by transient expression in mammalian cells, IRL Press at Oxford University Press, New York; Ausubel et al., 2000, Curr. Protocols in Molec. Biol). In contrast to the two-hybrid system technologies, it is the preferred technology to detect protein-protein interactions in the extracellular compartment. One version of this method is described here in detail for illustration only an not for limitation.

The expression cloning approach involves:

construction and/or purification of a probe, that will be used for screening

creation of a cDNA expression library from a suitable mRNA source

screening of the expression library and determination of the positive pools that express a ligand interacting with the probe

subcloning until a single clone bearing the cDNA of the ligand is found.

The screening of cDNA expression libraries in cultured mammalian cells is more laborious than the screening of a phage or plasmid library in bacterial cultures. It requires the amplification of the plasmid DNA

of the cDNA library clones in bacteria, its isolation and the subsequent transfection into mammalian cells for expression. Yet, mammalian cells, such as COS cells, are ideal hosts for said purpose. These cells are able to synthesize long translation products correctly from the cDNA template and to carry out the folding of the protein and its postranslational modifications in the extracellular compartment correctly. One possibility to design the probe for the screening of a cDNA expression library is fusion of the protein of interest with alkaline phosphatase (Flanagan and Leder, 1990, Cell 63, 185-194). Alkaline phosphatase (AP) has an intrinsic enzyme activity that can be used to trace the fusion protein with high sensitivity. A wide variety of substrates for AP allows quantitative assays in solutions or *in situ* detection. Since antibodies against AP are available, the immunological detection of AP-fusion protein is also possible.

For the production of the AP-tagged fusion protein, eucaryotic expression vectors, such as pcDNA3.1 (Invitrogen) are suitable. Because the interaction of a calsyntenin with its ligand may be perturbed, if the binding site is located near the region where the AP tag is fused, two constructs need to be generated for each fragment of a calsyntenin included in the screening program, one with the AP tag fused to the 3'-end and another with the AP tag fused to the 5'-end. The fusion proteins can be produced by transient expression in a suitable eucaryotic cell line, such as HEK 293T or COS. The well-expressed proteins are tested for example by Western blot analysis and quantified by an AP assay. As calsyntenin proteins are expressed in neurons throughout the brain, including the hippocampus, the probes may be tested for binding to cultures of dissociated hippocampal neurons. Testing of the probes with cultured neurons from other brain regions or with tissue slices is also possible.

To obtain a control for expression and binding experiments the sequence of secreted AP can be in-

serted in the same expression vector. Use of the endogenous secretion signal and the Kozak sequence of a calsyntenin or AP, together with the CMV promoter warrants the efficient translation of the fusion protein and its secretion into the culture medium.

The vector thus generated is transfected into the human embryonic kidney cell line, HEK293T, using the calcium phosphate transfection method or another transfection method, such as electroporation or lipofection. The conditioned medium is collected 3-4 days after transfection. The amount of the expressed protein is estimated by measuring the AP activity in the conditioned medium. A sample of conditioned medium is diluted in a buffer containing the soluble substrate of AP. The velocity of conversion of the substrate by AP is proportional to the activity of this enzyme. Since AP converts the substrate into a color product, it is possible to measure it spectrofotometrically. The expressed proteins are also tested by Western blot analysis using a polyclonal antibody against AP. Thus, the apparent molecular weight of the fusion protein may be determined as a control for its integrity.

Eucaryotic expression cloning requires a suitable RNA-source for generating a cDNA library. The mRNAs encoding the calsyntenin proteins are expressed predominantly in the neurons of the central nervous system, including the neurons of the hippocampus. Immunohistochemical staining of mouse brain sections confirmed the calsyntenin-1 expression in these brain regions also on the protein level. Therefore, it is plausible to assume that the putative ligand is expressed in the same regions where a calsyntenin protein is located. To test this hypothesis the dissociated hippocampal cell cultures were prepared and tested on binding of the probes. The binding assay is carried out according to the following protocol. The dissociated hippocampal neurons are shortly prefixed and incubated for 90 minutes with the buffered condi-

tioned medium containing the appropriate probe. Subsequently the cells are washed and subjected to a short pre-fixation. Endogenous heat-sensitive AP is then inactivated by incubation of cells at 65 °C for 2 hours. The AP inserted into the fusion proteins is heat-stable and remains active after this step. At the end the cells are incubated with the staining buffer containing a substrate that can be converted by AP into a colored precipitate. Therefore, the cells that express a molecule that binds the probe are stained blue.

The cDNA for the generation of a expression library can be generated from mRNA obtained from the brain of an adult mouse, or rat, or human by a standard technique (see Ausubel et al., 2000). Eucaryotic expression vectors for the transfection of the library into COS cells include for example pCDM8 (Aruffo and Seed, 1987, Proc. Natl. Acad. Sci. USA 84, 8573-8577) or, more recently, pcDNA31 (Invitrogen). For the generation of the library, various protocols have been successfully used (see e.g. Simmons, 1993). Immediately after generation, a cDNA expression library can be divided into approximately 200 pools with complexity 1000 - 1500 colony-forming units (cfu) per pool. Each of the pools is plated out for example in triplicate. 500-1.000 cfu are grown on each plate. After 36 hours, when the colonies have reached to diameter of 2-3 mm, the bacteria are washed from the plates with the medium. A part of the bacterial suspension can be mixed with glycerol and stored frozen as a back-up for subsequent subpooling. The rest of the suspension can be used for the isolation of the plasmid DNA. For example COS cells are transfected with the plasmid DNA of individual cDNA library pools and after 48 hours, when the cDNA fragments are expressed; they are tested on the probe binding. The efficiency of the transfection and quality of the staining reaction was always controlled by transfection of the cells with cDNA of neuropilin-1 and staining of these control cells with its known binding

partner semaphorin-III fused to AP. As a negative control mock transfected cells stained with both calsyntenin-AP and semaphorin-AP. All the cDNA library pools can be screened in triplicate.

All the positive pools can be subjected to the subpooling procedure. From the back-up of each positive pool 50 plates are plated, so that on each plate about 100 cfu are present. When the colonies become visible a replica is made. Both replica and original plate are incubated further till the colonies reach a diameter of about 2-3 mm. Then the bacteria are washed from the replica plates and the plasmid DNA is isolated. The original plates are stored at 4°C for the next round of subpooling. The COS cells are then transfected with the isolated DNA and after 2 days tested with the same probe.

Once a pathway gene has been identified and isolated, it may be further characterized as, for example, discussed herein.

The proteins identified as products of pathway genes may be used to modulate gene expression of a calsyntenin, as defined herein. Aternatively, the proteins identified as products of pathway genes may be used to modulte the proteolytic cleaveage of a calsyntenin and the resulting internalization of the transmembrane fragement of a calsyntenin with its calcium-binding cytoplasmic domain. Alternatively, the proteins identified as products of pathway genes may be used to modulate the influence of the calcium-binding domain of a calsyntenin in synaptic calcium signaling. Pathway genes may themselves be targets for modulation to in turn modulate calsyntenin protein function.

The compounds identified in the screen will demonstrate the ability to selectively modulate the activity of a calsyntenin protein as a modulator of synaptic calcium signaling. These compounds include, but are not limited to, small organic molecules that regulate the

proteolytic activity of the protease(s) that cleave(s) calsyntenin proteins in the extracellular part, molecules that bind the extracellular part of a calsyntenin protein and thereby modulate the susceptibility of a calsyntenin protein for proteolytic cleavage and/or internalization, molecules that bind to the transmembrane or cytoplasmic domain of a calsyntenin protein and modulate its affinity for calcium or its capacity of binding calcium, and molecules that bind to any part of a calsyntenin protein and modulate its interaction with macromolecular ligands, such as those defined herein as pathway proteins or pathway genes. These compounds also include, but are not limited to, nucleic acid encoding a calsyntenin protein and homologues, analogues, and deletions thereof, as well as antisense, ribozyme, triple helix, double-stranded RNA, antibody, and polypeptide molecules and small inorganic or organic molecules.

Any of the identified compounds can be administered to an animal host, including a human patient, by itself, or in pharmaceutical compositions where it is mixed with suitable carriers or excipient(s) at doses therapeutically effective to treat or ameliorate a variety of disorders, including those characterized by insufficient, aberrant, or excessive calsyntenin activity. A therapeutically effective dose further refers to that amount of the compound sufficient to result in amelioration of symptoms associated with such disorders. Techniques for formulation and administration of the compounds of the instant application may be found in "Remington's Pharmaceutical Sciences," Mack Publishing Co., Easton, Pa., latest edition.

A number of disorders may result from insufficient, aberrant, or excessive calsyntenin protein activity. In addition, several physiological states which may, from time to time be considered undesired, may also be associated with calsyntenin activity. By way of example, but not by way of limitation, such disorders and

physiological states which may be treated with the compounds of the invention include but are not limited to psychiatric disorders such as schizophrenia or depression, neurologic disorders such as Alzheimer's disease, stroke, and acute head injury, acute or chronical headache, hypertension, and myocardial infarction.

Other options may include direct delivery of enzyme which has been produced and purified by genetic means using the cloned gene. Other isoforms may exist and may be cloned utilizing a calsyntenin sequence. The compounds of the invention may be designed or administered for tissue specificity. If the compound comprises a nucleic acid molecule, including those comprising an expression vector, it may be linked to a regulatory sequence which is specific for the target tissue, such as the brain, kidney, heart, etc. by methods which are known in the art including those set forth in Hart, 1994, Ann. Oncol., 5 Suppl 4: 59-65; Dahler et al., 1994, Gene, 145: 305-310; DiMaio et al., 1994, Surgery, 116:205-213; Weichselbaum et al., Cancer Res., 54:4266-4269; Harris et al., 1994, Cancer, 74 (Suppl. 3):1021-1025; Rettinger et al., Proc. Nat'l. Acad. Sci. USA, 91:1460-1464; and Xu et al, Exp. Hematol., 22:223-230; Brigham et al., 1994, Prog. Clin. Biol. Res., 388:361-365. The compounds of the invention may be targeted to specific sites by direct injection to those sites. Compounds designed for use in the central nervous system should be able to cross the blood brain barrier or be suitable for administration by localized injection. In addition, the compounds of the invention which remain within the vascular system may be useful in the treatment of vascular inflammation which might arise as a result of arteriosclerosis, balloon angioplasty, catheterization, myocardial infarction, vascular occlusion, and vascular surgery. Such compounds which remain within the bloodstream may be prepared by methods well known in the art including those described more

fully in McIntire, 1994, Annals Biomed. Engineering, 22:2-13.

Pharmaceutical compositions suitable for use in the present invention include compositions wherein the active ingredients are contained in an effective amount to achieve its intended purpose. More specifically, a therapeutically effective amount means an amount effective to prevent development of or to alleviate the existing symptoms of the subject being treated. Determination of the effective amounts is well within the capability of those skilled in the art, especially in light of the detailed disclosure provided herein.

For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from cell culture assays. For example, a dose can be formulated in animal models to achieve a circulating concentration range that includes the IC50 (the dose where 50% of the cells show the desired effects) as determined in cell culture. Such information can be used to more accurately determine useful doses in humans.

A therapeutically effective dose refers to that amount of the compound that results in amelioration of symptoms or a prolongation of survival in a patient. Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., for determining the LD50 (the dose lethal to 50% of the population) and the ED50 (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio between LD50 and ED50. Compounds which exhibit high therapeutic indices are preferred. The data obtained from these cell culture assays and animal studies can be used in formulating a range of dosage for use in human. The dosage of such compounds lies preferably within a range of circulating concentra-

tions that include the ED50 with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. The exact formulation, route of administration and dosage can be chosen by the individual physician in view of the patient's condition. (See e.g. Fingl et al., 1975, in "The Pharmacological Basis of Therapeutics", Ch. 1 p1). Dosage amount and interval may be adjusted individually to provide plasma levels of the active moiety which are sufficient to maintain the desired effects.

In cases of local administration or selective uptake, the effective local concentration of the drug may not be related to plasma concentration.

The amount of composition administered will, of course, be dependent on the subject being treated, on the subject's weight, the severity of the affliction, the manner of administration and the judgement of the prescribing physician.

The pharmaceutical compositions of the present invention may be manufactured in a manner that is itself known, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping or lyophilizing processes.

Pharmaceutical compositions for use in accordance with the present invention thus may be formulated in conventional manner using one or more physiologically acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Proper formulation is dependent upon the route of administration chosen. For injection, the agents of the invention may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks's solution, Ringer's solution, or physiological saline buffer. For transmucosal administration, penetrants appropriate to

the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

For oral administration, the compounds can be formulated readily by combining the active compounds with pharmaceutically acceptable carriers well known in the art. Such carriers enable the compounds of the invention to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions and the like, for oral ingestion by a patient to be treated. Pharmaceutical preparations for oral use can be obtained solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are, in particular, fillers such as sugars, including lactose, sucrose, mannitol, or sorbitol; cellulose preparations such as, for example, maize starch, wheat starch, rice starch, potato starch, gelatin, gum tragacanth, methyl cellulose, hydroxypropylmethyl-cellulose, sodium carboxymethylcellulose, and/or polyvinylpyrrolidone (PVP). If desired, disintegrating agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, or alginic acid or a salt thereof such as sodium alginate.

Dragee cores are provided with suitable coatings. For this purpose, concentrated sugar solutions may be used, which may optionally contain gum arabic, talc, polyvinyl pyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for identification or to characterize different combinations of active compound doses.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a plasticizer, such as glycerol or sorbitol. The push-fit capsules can contain the active ingredients in admixture

with filler such as lactose, binders such as starches, and/or lubricants such as talc or magnesium stearate and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid paraffin, or liquid polyethylene glycols. In addition, stabilizers may be added. All formulations for oral administration should be in dosages suitable for such administration. For buccal administration,the compositions may take the form of tablets or lozenges formulated in conventional manner. For administration by inhalation, the compounds for use according to the present invention are conveniently delivered in the form of an aerosol spray presentation from pressurized packs or a nebulizer, with the use of a suitable propellant, e.g., dichlorodifluoromethane, trichlorofluoromethane, dichlorotetrafluoroethane, carbon dioxide or other suitable gas. In the case of a pressurized aerosol the dosage unit may be determined by providing a valve to deliver a metered amount. Capsules and cartridges of e.g. gelatin for use in an inhaler or insufflator may be formulated containing a powder mix of the compound and a suitable powder base such as lactose or starch.

The compounds may be formulated for parenteral administration by injection, e.g., by bolus injection or continuous infusion. Formulations for injection may be presented in unit dosage form, e.g., in ampoules or in multidose containers, with an added preservative. The compositions may take such forms as suspensions, solutions or emulsions in oily or aqueous vehicles, and may contain formulatory agents such as suspending, stabilizing and/or dispersing agents.

Pharmaceutical formulations for parenteral administration include aqueous solutions of the active compounds in water-soluble form. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic

solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Optionally, the suspension may also contain suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated solutions. Alternatively, the active ingredient may be in powder form for constitution with a suitable vehicle, e.g., sterile pyrogen-free water, before use. The compounds may also be formulated in rectal compositions such as suppositories or retention enemas, e.g., containing conventional suppository bases such as cocoa butter or other glycerides. In addition to the formulations described previously, the compounds may also be formulated as a depot preparation. Such long acting formulations may be administered by implantation (for example subcutaneously or intramuscularly) or by intramuscular injection. Thus, for example, the compounds may be formulated with suitable polymeric or hydrophobic materials (for example as an emulsion in an acceptable oil) or ion exchange resins, or as sparingly soluble derivatives, for example, as a sparingly soluble salt.

A pharmaceutical carrier for the hydrophobic compounds of the invention is a cosolvent system comprising benzyl alcohol, a nonpolar surfactant, a water-miscible organic polymer, and an aqueous phase. Naturally, the proportions of a co-solvent system may be varied considerably without destroying its solubility and toxicity characteristics. Furthermore, the identity of the co-solvent components may be varied.

Alternatively, other delivery systems for hydrophobic pharmaceutical compounds may be employed. Liposomes and emulsions are well known examples of delivery vehicles or carriers for hydrophobic drugs. Certain or-

ganic solvents such as dimethylsulfoxide also may be employed, although usually at the cost of greater toxicity. Additionally, the compounds may be delivered using a sustained-release system, such as semipermeable matrices of solid hydrophobic polymers containing the therapeutic agent. Various of sustained-release materials have been established and are well known by those skilled in the art. Sustained-release capsules may, depending on their chemical nature, release the compounds for a few weeks up to over 100 days. Depending on the chemical nature and the biological stability of the therapeutic reagent, additional strategies for protein stabilization may be employed.

The pharmaceutical compositions also may comprise suitable solid or gel phase carriers or excipients. Examples of such carriers or excipients include but are not limited to calcium carbonate, calcium phosphate, various sugars, starches, cellulose derivatives, gelatin, and polymers such as polyethylene glycols.

Many of the compounds of the invention may be provided as salts with pharmaceutically compatible counterions. Pharmaceutically compatible salts may be formed with many acids, including but not limited to hydrochloric, sulfuric, acetic, lactic, tartaric, malic, succinic, etc. Salts tend to be more soluble in aqueous or other protonic solvents that are the corresponding free base forms.

Suitable routes of administration may, for example, include oral, rectal, transmucosal, transdermal, or intestinal administration; parenteral delivery, including intramuscular, subcutaneous, intramedullary injections, as well as intrathecal, direct intraventricular, intravenous, intraperitoneal, intranasal, or intraocular injections.

Alternately, one may administer the compound in a local rather than systemic manner, for example, via injection of the compound directly into an affected area,

often in a depot or sustained release formulation. Furthermore, one may administer the drug in a targeted drug delivery system, for example, in a liposome coated with an antibody specific for affected cells. The liposomes will be targeted to and taken up selectively by the cells.

The compositions may, if desired, be presented in a pack or dispenser device which may contain one or more unit dosage forms containing the active ingredient. The pack may for example comprise metal or plastic foil, such as a blister pack. The pack or dispenser device may be accompanied by instructions for administration. Compositions comprising a compound of the invention formulated in a compatible pharmaceutical carrier may also be prepared, placed in an appropriate container, and labelled for treatment of an indicated condition. Suitable conditions indicated on the label may include treatment of a disease such as one characterized by insufficient, aberrant, or excessive calsyntenin-1 activity.

The just outlined uses of nucleic acid sequences and amino acid sequences as defined above has been shown in the scope of the present invention to be very suitable for protease involving disorders, in particular tPA involving diseases, and especially suitable for the treatment of stroke. For example in stroke treatment the calsyntenins or the calsyntenin derived proteins are suitable pharmaceuticals in acute treatment as well as in long-time treatment.

In an acute state, i.e. within the first few hours after a stroke, a presently preferred mode of application is the direct application of a high amount of a calsyntenin protein, preferably an intrathecal application, i.e. an injection directly into the cerebro-spinal fluids.

For the long term therapy of stroke, i.e. the restitution of damages, a preferred method is cell therapy.

For gene therapy and/or cell therapy a nucleic acid sequence coding for a calsyntenin protein (the expression a calsyntenin protein is considered as including alleles and mutants with protease inhibitor, at least tPA inhibitor activity) is introduced into a suitable vector allowing the expression of a calsyntenin gene in the addressed nerve cells or specific therapy cells. Such a vector suitable for gene therapy and allowing expression of the calsyntenin comprises the calsyntenin-1 encoding gene under the control of a nerve cell specific promoter.

For gene therapy suitable vectors are neurotrophic viruses that can be applied either directly or in transport cells.

Calsyntenin expressing cells can also be encapsulated so that they can be brought to the center of desired action by surgery treatment and with much reduced risk for incompatibility reactions. Such cells can be removed as soon as they are no longer needed or as soon as they have lost their activity and thus need replacement.

All the above described methods for the treatment of stroke are similarly applicable to other disorders induced by proteases, in particular tPA. Such disorders also comprise tumors such as those induced by tPA due to its effect on cell migration, but also tumors generally involving at least one protease in their growth, expansion, infiltration, metastasis and promotion of blood vessels or neoangiogenesis. Such proteases are preferably members of at least one of the following protease families:

- Serine Protease family such as tissue-type plasminogen activator (tPA), urokinase-type plasminogen activator (uPA), plasmin, thrombin, elastases, cathepsin G, neuropsin, neurotrypsin

- Matrix Metalloproteinases family such as collagenases, gelatinases, stromelysins, matrilysins,

- Cystein Proteases family such as cathepsin B and cathepsin D.

Besides of the above further described treatments, the present invention also provides for very useful diagnostic tools. By PCR and hybridization methods, as already mentioned above, genetic defects in the calsyntenin encoding protein can be determined. Such determination helps for the diagnosis of disorders the symptoms of which are already noticeable as well as for the determination of persons or groups of persons, such as families, with enhanced risk to develop such a disorder.

It is of course also possible to produce by synthetic or chemical methods proteins, peptides or nucleic acid sequences representing at least part of the sequences defined above and having the ability to mimic or to block, respectively, the biological activity of calsyntenin, in particular the calcium binding activity.

Furthermore, the characterization and isolation of a deficient gene or a deficient protein encoded by such a gene provides efficient tools for screening possible drugs to improve the health of patients suffering from disorders due to such defects.

In particular for the search of further disorders and drugs also transgenic animals are of great value.

This and further aspects of the present invention are now further illustrated by the following examples. It has, however to be understood that they are not at all intended to reduce the scope of the present invention. They are of mere illustrative purpose.

Example 1:

Screening for proteins released from the neurites of embryonic chicken spinal cord neurons identifies

a 115 kD protein as a proteolytic fragment of calsyntenin-1.

In a search for proteins released from neurites we cultivated dissociated spinal cord neurons in a compartmental cell culture system that provides separate access to neuronal cell bodies and neurites (Fig. 1C). The compartmental cell culture system was set up as described by Campenot (1979), Methods Enzymol, 58, 302-7. Dissociated cells from the ventral halves of spinal cords of E6 chicken embryos were cultivated in the center compartment of the compartmental culture system as described previously (Sonderegger et al. 1984, J.Cell. Biol. 98(1): 364-8). Six days after plating, when the side compartments had become densely populated by neurites (Fig. 1B), the newly synthesized proteins were metabolically labeled by adding fresh medium containing [^{35}S] methionine to the center compartment (Stoeckli et al., 1989, Eur. J. Biochem. 180(2): 249-58). After 40 hours, the conditioned media of both the center and the side compartments were harvested and subjected to two-dimensional gel electrophoresis (O'Farrell, J Biol Chem. 1975 May 25;250(10): 4007-21.) followed by fluorographic detection (Bonner and Laskey, Eur J Biochem. 1974 Jul 1;46(1):83-8.) of the newly synthesized proteins (Fig. 1D and E). As shown in Fig. 1D, the supernatant of the center compartment contained a relatively large number of proteins, whereas only four strong protein spots were found in the side compartment (Fig. 1E). Because proteins diffusing from the center to the side compartment did not reach more than 10 % of their concentration in the center compartment, we concluded that these four proteins had to be derived from the neurites of the side compartment (for a quantitative study with the same system see (Stoeckli et al., 1989, Eur. J. Biochem. 180(2): 249-58). One of them (Fig. 1E, arrow 1) was previously identified as neuroserpin, an axonally secreted serine protease inhibitor (Osterwalder et al., EMBO J. 1996; 15(12):2944-53.).

Based on molecular weight and pI, three other proteins were unknown. The protein with an apparent molecular weight of 115 kD and a pI of 5.9 to 6.3 (arrow 2 in Fig. 1E) was isolated and characterized as reported in the following examples. Purification, amino acid sequencing, and cDNA cloning (see examples 2 and 3) revealed that the 115 kD protein released from the neurites of embryonic chicken spinal cord neurons is a proteolytic fragment of a transmembrane protein. Because of its synaptic localization (see examples 10 and 11) and its capacity to bind calcium with its cytoplasmic domain (example 6), we termed it calsyntenin-1. For brevity and clarity, we will use the term calsyntenin-1 throughout the examples.

Example 2:

Purification and microsequencing of the 115 kD fragment released from the transmembrane-anchored calsyntenin-1 protein.

For the purification of the 115 kD fragment of calsyntenin-1 that is released from the neurites of spinal cord neurons in the compartmental culture system (see example 1), we used the conditioned medium of dissociated cultures of the ventral halves of spinal cords from E6 chickens. 6 x 106 cells from the ventral halves of the spinal cords of E6 chickens were cultivated in 60 mm collagen-coated culture dishes (porcine collagen, 25010 COL1, Corning, NY) and grown for 7 days with one change of medium. To harvest released proteins, the cells were washed twice with prewarmed MEM without supplements and grown for 2-3 days in serum-free medium with nutrient mixture N3 (Bottenstein and Sato, Proc. Natl. Acad. Sci. USA 1979, 76(1): 514-7), lacking BSA and transferrin. The conditioned medium was harvested, filtrated trough a 0.22 µm filter and stored at -20 °C.

For the purification of calsyntenin-1, the conditioned medium was dialyzed against buffer A (20 mM Tris-Cl, pH 8.0), degased, filtrated again through a 0.22

µm filter, and then loaded onto a 1 ml Mono Q anion exchange column (Pharmacia) at a flow rate of 1 ml/min. After washing the column with 10 volumes of buffer A, the proteins were eluted in a gradient from 0 % to 50 % of buffer B (1 M NaCl in 20 mM Tris-Cl, pH 8.0) within 20 ml. Fractions of 3 ml were collected and analyzed with 2-dimensional SDS-PAGE (O'Farrell, J Biol Chem. 1975 May 25;250(10): 4007-21.) followed by silver staining (Heukeshoven and Dernick, Electrophoresis. 1988;9(1):28-32.). Calsyntenin-1 was eluted between 300 and 450 mM NaCl.

For preparative separation by 2-dimensional SDS-PAGE, the fractions were pooled and concentrated either according to Wessel and Fluegge, (Anal. Biochem. (1984), 138:141-3)or by centrifugation through a porous membrane (Ultrafree-20, Milipore, Bedford, MA). Two-dimensional SDS-PAGE was carried out according to O'Farrell (1975), loading 3-4 concentrated fractions from the anion exchange column onto one gel. The ampholine solution for the isolelectric focusing step was composed of 1.6 % pharmalyte 5/8, 0.4 % pharmalyte 3/10, and 0.8 % pharmalyte 4/6 (all from Pharmacia). The pH range of the gels during isoelectric focusing was from pH 4.9 to 6.8. The second dimension was run on a 7.5 % SDS-PAGE gel (Lämmli, 1970).

After Coomassie blue staining, the protein spots with the gel coordinates of calsyntenin-1 were excised and processed by SDS-PAGE using the funnel-well concentration system (Lombard-Platet and Jalinot, Nucleic Acids Res. 1993, 21(17):3935-42). The funnel-well gel electrophoresis system devised by Lombard-Platet and Jalinot (1993) is a method for the concentration of protein from several gel pieces. Two spacers forming a funnel were adapted to the minigel system of Bio-Rad (Bio-Rad, Richmond, CA). Sealing was done with 20 % acrylamide. The running gel composed of 10 % acrylamide had a length of 1 cm, the stacking gel composed of 4 % acrylamide was 4-5 cm long.

Prior to loading into the funnel-well, the gel pieces containing calsyntenin-1 were destained from Coomassie blue in 20 % ethanol/5 % acetic acid and equilibrated for 3 h ar room temperature in sample buffer (4 % glycerol, 2.5 % SDS, 2.5 % β-mercaptoethanol in 25 mM Tris-Cl, pH 6.8). Up to 8 gel pieces containing calsyntenin-1 were carefully transferred into the funnel-well, overlayed with running buffer, and concentrated in the gel for 2 h at 50V. During the concentration, the current changed from 10 mA to 4 mA. The progress of the concentration could be followed by visual inspection of the protein front within the gel due to a schlieren effect. When the protein was concentrated in the middle of the running gel, the run was stopped for in-gel digestion or for transfer onto a PVDF membrane for direct amino acid sequencing. In a typical experiment, a protein band containing approximately 20 µg of calsyntenin-1 was obtained. After excision and destaining with 40 % n-propanol (LichroSolv grade, Merck), calsyntenin-1 was extracted with 0.2 M NH_4HCO_3, 50 % acetonitrile, and dried in a Speed Vac. For sequencing the N-terminus, concentrated calsyntenin-1 was electrotransferred from the gel onto a PVDF membrane (Immobilon P^{SQ}, Millipore). The calsyntenin-1-containing area on the PVDF membrane was localized by autoradiography and excised. The sequence was determined by Edman degradation on a protein sequencer (Model 477 A, Applied Biosystems, Inc.).

For sequencing of internal peptides, tryptic digestion was carried out within the gel pieces obtained by the funnel-well system (Jeno et al., Anal Biochem. 1995;224(1):75-82.). Calsyntenin-1 was processed as described above by 2-dimensional SDS-PAGE, stained with Coomassie blue, excised, and concentrated in the funnel-well system. The protein band of approximately 20 µg calsyntenin-1 was cut into small pieces. The gel pieces were destained with 40 % n-propanol (LichroSolv grade, Merck), extracted with 0.2 M NH_4HCO_3, 50 % acetonitrile, and dried

completely in a Speed Vac for 30 min. Tryptic digestion was carried out in digestion buffer (5 % acetonitrile in 100 mM Tris-Cl, pH 8.0), containing 1 µg trypsin (Promega, Madison, WI; 0.5 µg/µl in 1 mM HCl) and incubated at 37 °C for 18 h. The peptides were extracted with digestion buffer and with 80 % acetonitrile, 0.1 % trifluoroacetic acid. The pooled extracts were evaporated to obtain the injection volume of the reversed-phase HPLC column (30 - 50 µl). The peptides were separated in a reversed-phase HPLC column (Vydac C8, 5 µm particle sized, 1mm (i.d.) x 250 mm; Vydac, Hesperia, CA) connected to a mass spectrometer (API-III, PE Sciex, Thornhill, Ontario). Solvent A was 0.1 % trifluoroacetic acid, solvent B was 80 % acetonitrile containing 0.09 % trifluoroacetic acid. The elution program used was: 5 % solvent B for 5 min; 5 % to 60 % solvent B during 60 min at a flow rate of 50 µl/min. The effluent was monitored at 215 nm. 90 % of the eluted volume were collected and 10 % injected online into the mass spectrometer (API-III, PE Sciex, Thornhill, Ontario. The chromatograms were analyzed and single peptide fractions chosen for sequencing. Sequence analysis was performed on a Model G 1005 protein sequencer (Hewlett-Packard, Camas, WA), according to the manufacturers protocols.

By this method, the amino acid sequences of the N-terminus and seven internal peptides of the 115 kD fragment of calsyntenin-1 released from the cultures of spinal cord neurons were determined. The following sequences were found:

N-terminal sequence (single letter code for amino acids):

ARVNKHKPWIETTY (Seq. Id. No. 7)

Internal peptides:

Peptide Number	Amino acid sequence
1.	HKPWIETTYHGIVTENDNTVLLDP (Seq. Id. No. 8)
2.	VEAVDA (Seq. Id. No. 9)

3. IEYEPGTGSLALFPSMR (Seq. Id. No. 10)
4. IPDGVVT (Seq. Id. No. 11)
5. TYKPAEFHW (Seq. Id. No. 12)
6. EGLDLQIADGV (Seq. Id. No. 13)
7. GIEMSSSNLGMIITGVDTMASYEEVLHL (Seq. Id. No. 14)

The sequence of one internal peptide (Peptide Number 1) overlapped with the N-terminal sequence and, thus, generated an extension of the N-terminal sequence to a length of 29 amino acids, with the following sequence:

ARVNKHKPWIETTYHGIVTENDNTVLLDP

Example 3:

Cloning and sequencing of the calsyntenin-1 cDNA of the chicken

The amino acid sequences of the N-terminus and the internal peptides were used to design degenerated primers for RT-PCR using total RNA from E14 chicken brain as template. Total RNA was prepared from E14 chicken brain and from P10 mouse cerebellum (Chomczynski and Sacchi, Anal Biochem. 1987;162(1):156-9.). Oligo(dT)- and random-primed cDNA was produced using M-MLV reverse transcriptase (Promega). For PCR, degenerated primers corresponding to the amino terminus and four internal peptides were synthesized. (sense primers: 5'-GTIAAMAAGCAYAAGCCITGGAT-3' (Seq. Id. No. 15)and 5'-CATGGIATHGTIACIGAGAATGATAA-3' (Seq. Id. No. 16); antisense primers: 5'-CCIGTICCIGGCTCATACTCDAT-3'(Seq. Id. No. 17) , 5'-GTATCIACICCITADATDATCATICC-3'(Seq. Id. No. 18) , 5'-ACICCATCIGCDATCTGIAAATC-3' (Seq. Id. No. 19)and 5'-GCATCAAACTCIGCCTCCTTATAAAA-3' (Seq. Id. No. 20). PCR was performed using Taq DNA polymerase (Promega). PCR fragments were sequenced and used for screening cDNA libraries. Approximately 2.5 x 10^6 plaques of an oligo(dT)-primed E14 chicken brain cDNA library (Zuellig et al., 1992; Eur. J. Biochem 204(2):453-63), an oligo(dT)-primed

P20 mouse brain cDNA library (Stratagene), an oligo(dT)- and random-primed E15 mouse brain cDNA library (Clontech), and an oligo(dT)- and random-primed fetal human brain cDNA library (Clontech), respectively, were screened by hybridization with the corresponding radiolabeled PCR fragments under high stringency conditions (Sambrook et al., 1989; Molecular Cloning. A Laboratory Manual). Positive clones were further characterized.

The longest PCR product had a length of 2.2 kb. It contained a single open reading frame (ORF) that encoded all previously determined amino acid sequences (Fig. 2). Screening an oligo(dT)-primed E14 chicken brain cDNA-library (Zuellig et al., 1992; Eur. J. Biochem 204(2):453-63) with this fragment as a probe yielded clones containing additional 3´ sequence of the ORF and the 3´ untranslated region. The composite cDNA contained an ORF of 2850 nt (starting from the amino-terminus of the purified protein). The hydropathy plot provided evidence for a single transmembrane segment of 19 amino acids close to the C-terminus (Fig. 2). Therefore, we concluded that the mature protein was composed of an extracellular N-terminal moiety of 831 amino acids, a transmembrane segment of 19 amino acids, and a cytoplasmic moiety of 100 amino acids. Based on the presumed structural characteristics as a type I transmembrane protein, the 115 kD protein isolated from the supernatant of E6 spinal cord cultures represents the proteolytically cleaved N-terminal fragment of the full-length transmembrane protein. The exact location of the cleavage site within the sequence of full-length calsyntenin-1 remains to be determined. Based on the location of the tryptic peptides (boxed in gray in Fig. 2), the released fragment isolated from the culture supernatant must have a length of at least 750 amino acids (as counted from the N-terminus of the mature protein).

Thirty-eight of the 100 amino acids of the cytoplasmic segment of calsyntenin-1 are acidic (see Ta-

ble 1). In the most acidic middle part, 18 out of 20 residues are acidic and the flanking sequences are enriched in acidic residues as well. Similarly acidic segments are characteristic for calsequestrin, calreticulin, and protein disulfide isomerase (Fliegel et al., J. Biol. Chem. 1989; 264(36):21522-28). These proteins are essential for the storage of Ca^{2+} in the sarcoplasmic reticulum of skeletal muscle cells and the endoplasmic reticulum of nonmuscle cells, due to their capacity to bind large numbers of Ca^{2+} ions with low affinity (Baksh and Michalak, J Biol Chem. 1991; 266(32):21458-65; Ohnishi and Reithmeier, iochemistry. 1987; 26(23):7458-65.).

Example 4:

Cloning and sequencing of the calsyntenin-1 cDNA of the human and the mouse: Species homologues of calsyntenin-1 in vertebrates exhibit a high degree of structural conservation

The major part of the cDNA of human calsyntenin-1 was found by searching the THC (Tentative human consensus sequence) database with the THC Blast program. Seven THCs (THC176438, THC178825, THC195843, THC200424, THC192325, THC211114, and THC211115) with homology to the cDNA of chicken calsyntenin-1 were identified and used to compose a partial sequence of the human cDNA lacking a segment of the 5´ end and two internal segments. The gaps were closed by RT-PCR. The putative translation start codon and a segment of 5´ UTR sequence were found by screening a human brain cDNA library. Thus, we obtained a human cDNA sequence with an ORF of 2943 bp that was 100 % identical with KIAA0911, a cDNA resulting from a screen for brain-specific proteins that was not further characterized (Nagase et al., DNA Res. 1998; 5 (6), 355-364).

The cDNA of mouse calsyntenin-1 was obtained by RT-PCR and subsequent screening of brain cDNA libraries. Based on the sequence of overlapping clones, a sin-

gle ORF of 2937 nt, encoding a peptide of 979 amino acids was defined (Fig. 2).

The sequences of human and mouse calsyntenin-1 starting with the amino acid 29 correspond to the sequence of the N-terminal peptide of chicken calsyntenin-1 (Fig. 2). The deduced amino acid sequences of the human and the mouse orthologs had an identity of 86.4 % and 84.7 %, respectively to chicken calsyntenin-1 (Table I). The amino acid sequence identity of human and mouse calsyntenin-1 was 92 %.

These results revealed that the species homologues of calsyntenin-1 in three vertebrate species, namely human, mouse, and chicken, exhibit a high degree of structural conservation. Based on the high structural conservation, a high degree of functional conservation among the species orthologs of calsyntenin-1 of different vertebrate species, such as human, mouse, rat, and chicken, can be expected.

Example 5:

Database searches for calsyntenin-1-related genes: Related sequences found in databases suggest the existence of calsyntenin-like genes in D. melanogaster and C. elegans

Genes with a structural relationship to vertebrate calsyntenin-1 were also found in the databases for D. melanogaster and C. elegans. In the Genome Annotation Database of Drosophila (GadFly), a single calsyntenin-1-like gene was found (acc. Nr. GC11059). Based on six overlapping ESTs (GM09293, HL03914, LD07408, LD 11689, GM10465, LD06216) we have determined the sequence of the corresponding cDNA (acc. Nr. AJ289018). The deduced protein exhibits an amino acid sequence identity of approximately 35 % with vertebrate calsyntenin-1. A further calsyntenin-1-related gene, B0034.3 with Accession. Nr. AAC38816, was found in C.elegans (see Table I).

These results indicate that calsyntenin-1-like genes are also found in invertebrates. Thus, calsyntenin-1 represents an evolutionarily ancient gene that has been well conserved throughout evolution.

Example 6:

Calcium-binding studies with the cytoplasmic segment of calsyntenin-1: The cytoplasmic segment of calsyntenin-1 binds calcium ions

The clustered occurrence of acidic amino acids is a typical trait of high-capacity, low-affinity Ca^{2+}-binding proteins found in vesicular Ca^{2+} stores, such as calsequestrin (Yano and Zarain-Herzberg, Mol Cell Biochem. 1994; 135(1):61-70.) and calreticulin (Krause and Michalak, Cell. 1997; 88(4):439-43). To test for the Ca^{2+}-binding capacity of calsyntenin-1, we generated a fusion protein of its cytoplasmic segment with the bacterial protein intein. A fusion protein of the cytoplasmic segment of mouse calsyntenin-1 and an N-terminal intein tag was expressed in bacteria using the Impact-CN system (New England Biolabs Inc.).

The cDNA of the cytoplasmic segment of calsyntenin-1 was amplified by PCR before it was inserted in frame in the multiple cloning region (MCS) of the pTYB11 vector (New England BioLabs, Inc.). The PCR was performed using the proofreading polymerase Pwo (Roche), the complete mouse cDNA of calsyntenin-1 as template and the primers LV38Fmax3 (5'-GGGGAACAGAAGAGCTGCACATCAGCGAACG-3') (Seq. Id. No. 21) and LV39Bmax3 (5'-CCCCCTCGAGTTAGTAGCTGAGTGTGGAG-3') (Seq. Id. No. 22). The PCR fragment was cloned into the MCS of pTYB11 using restriction sites SapI and XhoI. After ligation the plasmid was used to transform competent E. coli strain BL21DE. A single colony containing the correct plasmid was used for protein expression. One liter LB medium containing 100 µg/ml ampicillin was inoculated with a fresh colony. The culture was incubated in an air shaker at

37°C until the OD_{260} reached 0.6. Afterwards the culture was transferred to a 22°C shaker. The protein expression was induced by adding 0.5mM Isopropyl-B-D-Thiogalactoside (IPTG). After 6 h, the cells were spun down at 5000xg for 10 minutes at 4°C. The cell pellet was resuspended in 20 ml of cell lysis buffer (20 mM Tris; 500 mM NaCl; pH 8.0) and the cells were broken by sonication. To obtain a clarified cell extract the crude cell extract was centrifuged at 12000xg for 30 minutes. To purify the calsyntenin-1/intein fusion protein the clarified cell extract was loaded onto a chitin column and washed with a high flow rate (2 ml/min) and stringent wash conditions (1 M NaCl). The fusion protein was eluted with 3x SDS-PAGE sample buffer (187.5 mM Tris-HCl pH6.8; 6% SDS; 30% glycerol and 0.03% bromphenolblue) and incubation for 3 minutes at 99°C.

As a control maltose-binding protein (pMYB5 control plasmid; New England BioLabs, Inc.) fused to the intein tag was used.

The Ca^{2+}-binding assay was performed as described previously (Maruyama et al., J Biochem 1984; 95(2):511-9.). For SDS-PAGE, 5 µg of purified fusion protein were heated at 100°C for 5 min prior to loading on 10% polyacrylamide gels. The electrotransfer onto a nitrocellulose membrane (Schleicher & Schuell, Dassel, Germany) was performed at a constant voltage of 100 V for 1h at 4 °C, using a solution containing 20% methanol, 0.025 M Tris-Cl, and 0.129 M glycine (pH8.5) as the electrode buffer. After transfer, the membrane was soaked in a solution containing 60 mM KCl, 5mM $MgCl_2$, and 10mM imidazole-HCl, pH6.8, and the buffer was exchanged several times in an hour. Then, the membrane was incubated in the same buffer containing 0.5 µM, or 1 µM, or 5 µM, respectively, of $^{45}Ca^{2+}$ (28mCi/mg calcium , Amersham, Buckinghamshire, UK) for 10 min. The membrane was rinsed with distilled water for 2 minutes. Excess water was absorbed with Whatman No. 1 filter paper and the membrane was

dried at room temperature. For autoradiography, the blots were analyzed using a PhosphorImager (Molecular Dynamics, Sunnyvale, CA).

As shown in Fig. 3, a dose-dependent $^{45}Ca^{2+}$ signal overlapping with the calsyntenin-1/intein fusion protein was found when the nitrocellulose filters were incubated with Ca^{2+} concentrations of 0.5 μM, 1 μM, and 5 μM. No Ca^{2+} binding was observed with a fusion protein composed of the bacterial maltose-binding protein and intein (MBP/intein in Fig. 3).

With the calsyntenin-1/intein fusion protein, we observed an extensive precipitation when $CaCl_2$ was added. The addition of 50 mM or more Ca^{2+} caused the soluble calsyntenin-1/intein, but not MBP/intein, to precipitate (not shown). The Ca^{2+}-induced precipitation of calsyntenin-1/intein was prevented by the addition of an equimolar concentration of EDTA. Because precipitation was not observed at Ca^{2+} concentrations below 50 mM, we concluded that cross-bridging between cytoplasmic domains of calsyntenin-1 involves low-affinity binding of Ca^{2+} to sites which are clearly distinct from the high-affinity sites detected by $^{45}Ca^{2+}$ binding on the nitrocellulose membranes. With respect to the low-affinity binding capacity for Ca^{2+}, the cytoplasmic domain of calsyntenin-1 exhibits striking similarities with calsequestrin, the major calcium-binding protein of the sarcoplasmic reticulum of striated muscle cells. Calsequestrin exhibits a similar clustering of acidic residues cumulating in a contiguous stretch of 14 acidic residues. The Ca^{2+}-binding capacity of peptides with contiguous acidic residues has been linked to a general cation-binding capacity rather than specific Ca^{2+} sites. A comparison of proteins with such acidic stretches suggested that the Ca^{2+}-binding capacity was proportional to the content of acidic residues (Lucero et al., J Biol Chem. 1994; 269(37):23112-9). X-ray crystallography suggested that the low-affinity binding of Ca^{2+} occured via intercalation of Ca^{2+} between the

acidic C-terminal segments of calsequestrin dimers (Wang et al., Nat Struct Biol. 1998; 5(6):476-83). Similarly, calsyntenin-1 may bind Ca^{2+} by intercalation between its cytoplasmic moieties which are held in an ordered parallel orientation by transmembrane anchorage.

In summary, these results indicate that the cytoplasmic domain of calsyntenin-1 exhibits both high-affinity and low-affinity binding of Ca^{2+}.

Example 7:

Northern blot analyses of the tissue distribution of calsyntenin-1 mRNA: The brain is the tissue with the highest expression level of calsyntenin-1 mRNA.

In order to obtain information on the expression pattern of calsyntenin-1 in different human tissues, a human multiple-tissue Northern blot (Clontech) was hybridized with a 2.8 kb cDNA fragment of human calsyntenin-1 labeled with [α-^{32}P] dCTP (Amersham) using the Prime-it II random primer labeling kit (Stratagene). Hybridization was performed at 42 °C overnight and the hybridization signals were analyzed with a PhosphorImager (Molecular Dynamics). Northern blot analysis of poly(A)-enriched RNA from adult human tissues revealed a single species of calsyntenin-1 mRNA of approximately 5kb (Fig. 4C). The highest expression of calsyntenin-1 mRNA was observed in brain. Low signals were detected in heart, placenta, skeletal muscle, and kidney. No transcript was found in lung and liver.

Therefore, these results clearly demonstrate that the highest expression levels of calsyntenin-1 mRNA is found in the brain.

Example 8:

In situ hybridization analyses of the tissue distribution of calsyntenin-1 mRNA: Calsyntenin-1 is predominantly expressed in neurons

In order to determine the expression of calsyntenin-1 in the brain at cellular resolution, *in situ* hybridization was performed as described previously (Schaeren-Wiemers and Gerfin-Moser, Histochemistry 1993; 100(6): 431-40). In situ hybridization on cryosections from a E18 mouse revealed a strong cellular expression of calsyntenin-1 mRNA in the gray matter of the central and the peripheral nervous system (Fig. 4A). In a saggital section of an E18 mouse the following regions were labeled (Fig. 4A), the neocortex (nc), the hippocambal formation (hi), the caudate putamen (cpu), the thalamus (th), the hypothalamus (hyth), the cerebellum (ce), the pons (po), the trigeminal ganglion (tg), the dorsal root ganglia (drg), the olfactory epithelium (oe), the submandibular gland (sg) and the intestine (in). No calsyntenin-1 mRNA was detected in non-neural tissues, except in the submandibular gland. Control sections processed with the sense probe, showed no staining (). In the adult mouse, calsyntenin-1 mRNA was abundant in all areas of the gray matter (Fig. 4B). In the white matter, such as the corpus callosum (cc) no calsyntenin-1 expression was found. Inspection at higher magnification indicated a neuronal expression pattern in all areas of the CNS and the PNS. Most, if not all neurons, expressed calsyntenin-1 mRNA, yet considerable differences in the expression level were found. Northern blot analysis of calsyntenin-1 mRNA in adult human tissues is shown in Figure 4C. Two µg of purified polyA$^+$ RNA per lane form heart (He), brain (Br), placenta (Pl), lung (Lu), liver (Li), skeletal muscle (Sm), and kidney (Ki) were analysed with radiolabeled cDNA fragments of human calsyntenin-1. The molecular size scale is in kb. In Figure 4D is shown a Western blot analysis of chicken calsyntenin-1 protein. 150 µg of tissue extract from adult chicken brain (Br), heart (He), liver (Li), testis (Te), chicken cerebrospinal fluid (CSF), and human cerebrospinal fluid (hCSF) were subjected to SDS-PAGE and immunoblotting using polyclonal

antibodies R63 (left panel) and R71 (right panel) against calsyntenin-1. The molecular weight scale is in kD. Figure 4E shows a schematic drawing indicating the proteolytic cleavage site (arrow) on the calsyntenin-1 protein and the location of the recombinant peptide segments used for raising the R63 (shadowed) and the R71 (hatched) antibodies in the complete sequence of mature calsyntenin-1. Note that antibody R63 recognizes both the full-length form of calsyntenin-1 and the N-terminal cleavage product. In contrast, antibody R71 recognizes the transmembrane stump generated by the proteolytic cleavage of calsyntenin-1. The transmembrane domain (TM) is marked in black. Scale bars: (A), 2.5 mm; (B), 1.0 mm.

Example 9:

Characterization of calsyntenin-1 protein and its cleavage products: Calsyntenin-1 protein occurs as a full-length transmembrane protein, a membrane-bound C-terminal cleavage product, and a soluble N-terminal cleavage product

To analyze the tissue distribution of full-length calsyntenin-1 and its cleavage products two antibodies, termed antibody R63 and R71, respectively, were raised in rabbits. The immunogen for the R63 antiserum consisted of a 267 amino acid peptide starting at the N-terminus of chicken calsyntenin-1. The immunogen for the R71 antiserum consisted of an 87 amino acid peptide located immediately outside of the transmembrane segment of chicken calsyntenin-1. Both fragments were expressed with a His-tag in bacteria and purified using a NiNTA column (Qiagen).

Production of the R63 antigen:

The cDNA fragment of the 267 amino acids long peptide located at the N-terminus of chicken calsyntenin-1 was amplified by PCR before it was inserted in frame in the pTFT74 vector. The PCR was performed using the proofreading polymerase pfu (Stratagene), the cDNA of chicken

calsyntenin-1as template and the primers LV31Fchax3 (5'-GGGCCATGGCTCGTGTTAACAAGCATAAGCCCTGGATTG-3')(Seq. Id. No. 23) and LV32Bchax3 (5'-CCCAAGCTTAGTGGTGGTGGTGATGGT-GTGGTTCATCACATGTGTCC-3')(Seq. Id. No. 24). The PCR fragment was cloned into the pTFT74 vector using restriction sites NcoI and HindIII. After ligation the plasmid was transformed into competent E. coli strain BL21DE. A single colony containing the correct plasmide was used for protein expression. 1 liter LB medium containing 100 µg/ml ampicillin was inoculated with a fresh colony. The culture was incubated in an air shaker at 37°C until the OD_{260} reached 0.6. Afterwards the culture was transferred to a 22°C shaker. The protein expression was induced by adding 0.5mM IPTG. After 6 h the cells were spun down at 5000xg for 10 minutes at 4°C. The cell pellet was resuspended in 20 ml of cell lysis buffer (50 mM Tris; pH 8.0) and the cells were broken by sonication. To obtain a clarified cell extract the crude cell extract was centrifuged at 12000xg for 30 minutes. The fusion protein was purified using a NiNTA column (Qiagen) according to the instruction manual. The R63 antigen generated in this way is shown below (single letter code for amino acids; capital letters indicate amino acids found in calsyntenin-1; small letters indicate the initial methionine and the histidine-tag, respectively.

Protein sequence of R63 antigen (Seq. Id. No. 25)

ARVNKHKPW IETTYHGIVT ENDNTVLLDP PLIALDKDAP LRFAESFEVT VTKEGEICGF KIHGQNVPFE AVVVDKSTGE GIIRSKEKLD CELQKDYTFT IQAYDCGKGP DGANAKKSHK ATVHIQVNDV NEYSPVFKEK SYKATVIEGK RYDNILKVEA VDADCSPQFS QICNYEIVTP DVPFAIDKDG YIKNTEKLSY GKEHQYKLTV TAYDCGKKRA AEDVLVKISI KPTCKPGWQG WSKRIEYEPG TGSLALFPSM RLETCDEP

Production of the R71 antigen:

The cDNA fragment of the 87 amino acid long peptide used as antigen for generation of R71 antibody was amplified by PCR before it was inserted in frame in the pTFT74 vector. The PCR was performed using the proof-reading polymerase pfu (Stratagen), the cDNA of chicken calsyntenin-1as template and the primers MS1Fchax3 (5'-GGGCCATGATACGCTACAGAAACTGGCAC-3')(Seq. Id. No. 26) and MS2Bchax3 (5'-CCCAAGCTTAGTGGTGGTGGTGATGGTGAGTGGC-TGTACTTGGAACAAC-3')(Seq. Id. No. 27). The PCR fragment was cloned into the pTFT74 vector using restriction sites NcoI and HindIII. After ligation the plasmid was transformed into competent E. coli strain BL21DE. A single colony containing the correct plasmide was used for protein expression. 1 liter LB medium containing 100 µg/ml ampicillin was inoculated with a fresh colony. The culture was incubated in an air shaker at 37°C until the OD_{260} reached 0.6. Afterwards the culture was transferred to a 22°C shaker. The protein expression was induced by adding 0.5mM IPTG. After 6 h the cells were spun down at 5000xg for 10 minutes at 4°C. The cell pellet was resuspended in 20 ml of cell lysis buffer (50 mM Tris; pH 8.0) and the cells were broken by sonication. To obtain a clarified cell extract the crude cell extract was centrifuged at 12000xg for 30 minutes. The fusion protein was purified using a NiNTA column (Qiagen) according to instruction manual. The R71 antigen generated in this way is shown below (single letter code for amino acids; capital letters indicate amino acids found in calsyntenin-1; small letters indicate the initial methionine and the histidine-tag, respectively.

Protein sequence of R71 antigen (Seq. Id. No. 28)

mIRYRNWHTV SLFDRKFKLV CSELNGRYVS NEFKVEVNVI HTANPIEHAN HIAAQPQFVH PVHHTFVDLS GHNLANPHPF SVVPSTATgh hhhhh

The antisera against the proteins (antigen R63 and antigen R71) were raised in rabbits by injection of 50 µg protein in phosphate-buffered saline with complete Freund's adjuvans for the first injection and with incomplete Freund's adjuvans for the booster injections. The anti-calsyntenin-1 antibodies were affinity purified from the immuneserum by a passage over a protein-G column followed by an antigen-conjugated column.

Antibody R63, raised against the N-terminal 267 amino acids of the mature protein, detects both full-length calsyntenin-1 as well as the N-terminal part resulting from the proteolytic cleavage (Fig. 4E). Antibody R71, raised against a segment of 87 amino acids located adjacent to the transmembrane domain, detects the transmembrane stump generated by the proteolytic cleavage. In Western blots, calsyntenin-1 immunoreactive bands were found exclusively in brain extracts and in the cerebrospinal fluid (CSF; Fig. 4D). Extracts of all the other tissues that were tested, including heart, liver, testes (Fig. 4D), as well as kidney, lung, and spleen (not shown) did not exhibit calsyntenin-1 immunoreactivity. In brain extracts of adult chickens, two bands with apparent MWs of 150 and 115 kD were found with antibody R63, whereas a single band at 33 kD was detected with antibody R71 (Fig. 4D). The 115 kD band comigrated with the protein initially identified with the compartmental culture system as a released protein of the axo-dendritic compartment of spinal cord neurons. The 150 kD band represents most likely the full-length form of calsyntenin-1, based on the estimated size of the released fragment and the length of the transmembrane and cytoplasmic segments. In the cerebrospinal fluid, antibody R63 recognized only a single band of 115 kD, that corresponds to the soluble N-terminal cleavage product of calsyntenin-1. In contrast to brain extract, CSF did not contain full-length calsyntenin-1 or the transmembrane stump. Taken together, these results indicate that full-length calsyntenin-1 and its

cleavage products coexist in brain tissue. The N-terminal 115 kD fragment of calsyntenin-1 that is solubilized after proteolytic cleavage is also found in the CSF.

An investigation of human CSF indicated that human calsyntenin-1 is cleaved in the extracellular moiety the same way as calsyntenin-1 of other vertebrates. As demonstrated in Fig. 4E, a prominent calsyntenin-1-immunoreactive band with the same molecular weight as the N-terminal cleavage fragment of calsyntenin-1 found in mouse and chicken. These results indicate that the proteolytic cleavage of calsyntenin-1 in the extracellular domain is a characteristic that was conserved during evolution.

Example 10:

Studies of the subcellular localization of calsyntenin-1 protein at the light microscopic level: Cell surface-bound calsyntenin-1 is colocalized with established synaptic marker proteins

Immunoperoxidase staining of tissue sections of the hippocampus (Fig. 5A) and the cerebral cortex (not shown) revealed that calsyntenin-1 was abundant in synapse-rich regions. At higher magnification, a punctate appearance of the immunostaining in the neuropil (insert of Fig. 5A) was found, suggesting a synaptic localization of calsyntenin-1.

A detailed study of the subcellular location of cell surface-associated calsyntenin-1 was performed by immunofluorescence colocalization in cultures of dissociated hippocampal neurons. Cell suspensions of hippocampi dissected from brains of E17 mice were prepared by digestion with trypsin (0.25 % for 10 min at 37°C) and trituration using a blue Gilson tip. Cells were then plated onto acid-washed, poly-L-lysine-treated glass coverslips or poly-L-lysine-treated plastic dishes in DMEM supplemented with B27 (Gibco/Life Technologies), 0.25 mg/ml Albumax (Gibco/Life Technologies), 2 mM glutamine, and 0.1

M sodium pyruvate. Cultures were maintained for up to 4 weeks in a humidified incubator with 5 % CO_2 at 37°C.

Cells were fixed in 4 % paraformaldehyde and 4 % sucrose in PBS for 30 min at 37°C. After rinsing with PBS, cells were preincubated in 10 % fetal calf serum and 0.1 % glycine in PBS at room temperature for 1 h before incubation with the primary antibody in 3 % fetal calf serum in PBS at 4°C for 24-48 h. For the double-labeling experiments, primary antibodies were incubated together. Cells were washed for at least 30 min in three changes of PBS. For secondary antibodies FITC-conjugated goat anti-rabbit IgG (Cappel) and Cy3-conjugated donkey anti-mouse IgG (Jackson Immuno Research Laboratories, Inc.) were used. For stainings with anti-GluR2, the cells were permeabilized with 0.1 % saponin.

Established synaptic markers, such as synaptophysin, the α_2 subunit of the $GABA_A$ receptor, and the GluR2 subunit of the AMPA receptor were used as markers for presynaptic terminals and postsynaptic membranes, respectively. The antibody against the $GABA_A$ receptor subunit $\alpha 2$ was provided by Jean-Marc Fritschy. The antibodies against synaptophysin, PSD95, GluR1 and GluR2 were from Roche, Pharmingen, and Chemicon, respectively. As demonstrated in Fig. 5, B-D, calsyntenin-1 immunoreactivity exhibited a patchy pattern along neurite bundles. A very similar staining pattern was found with the antibodies against synaptophysin (Fig. 5B) and the $GABA_A$ receptor (Fig. 5C). In the overlay, the majority of the large areas labeled with antibodies against synaptophysin and the $GABA_A$ receptor were at least partially superposed, with the calsyntenin-1 immunoreactivity. With a commercially available antibody againt a cytoplasmic epitope of GluR2, which required permeablization of the cells and, therefore, stained both surface-exposed and internal AMPA receptors (Fig. 5D), large immunoreactive patches were found in close proximity to and sometimes partially overlapping with patches of calsyntenin-1 immunoreactivity.

Together, these results demonstrate the synaptic localization of calsyntenin-1.

Example 11:

Studies of the subcellular localization of calsyntenin-1 protein by immuno-electron microscopy: Full-length calsyntenin-1 is a component of the postsynaptic membrane.

To reveal the subcellular localization of calsyntenin-1 in CNS neurons, we have used preembedding and postembedding immuno-electronmicroscopy. To prepare brain tissue for immuno-EM, 8 adult Wistar and OFA line rats (200-250 g) of both sexes were deeply anaesthetized with metiofane (methoxyflurane, Pitman-Moore Inc., USA) and perfused through the ascending aorta for 15-25 min first with 0.9% saline for 1 min followed by fixative containing 3.5-4% paraformaldehyde, 0.015-0.05% glutaraldehyde, and 0.2% picric acid made up in 0.1M phosphate buffer pH 7.4. Then brains were removed from the skull into cold PB and either 70 µm (6 rats used in preembedding immunocytochemistry) or 500 µm thick coronal sections (2 rats used for freeze substitution) were cut on a vibratome.

For preembedding immunocytochemistry, the sections were cryoprotected in 30% sucrose, quickly frozen in liquid nitrogen and thawed in PB. After preincubation in 20% normal goat serum (NGS; Vector Labs, USA), sections were incubated in primary antibody made up in 0.05 mM Tris buffered saline pH 7.4 (TBS) containing 2% BSA and 2% NGS at 4°C for 2 days. For immunogold method, sections were incubated overnight in 1:40 goat anti-rabbit IgG coupled to 1.4 nm gold (Nanoprobes Inc. Stony Brook, NY), postfixed in 1% glutaraldehyde in PBS followed by silver enhancement of the gold particles with an HQ Silver kit (Nanoprobes Inc). For peroxidase reaction, sections were incubated for 4 h at RT in biotinylated goat anti-rabbit IgG (Vector Labs) diluted 1:200 in TBS

containing 1% NGS followed by 2 h incubation in avidin-biotin-peroxidase complex (ABC kit; Vector Labs) diluted 1:100 in TBS. Antigenic sites were revealed using standard 3,3'-diaminobenzidine tetrahydrochloride histostaining procedure (0.05% DAB and 0.01% H_2O_2 in TB pH 7.6). The gold-silver and peroxidase reacted sections were post-fixed in 1% osmium tetroxide in PB, stained with 2% uranyl acetate, dehydrated in graded series in ethanol and flat-embedded on glass slides in Durcupan ACM resin (Fluka) for electron microscopy.

For postembedding immunocytochemistry on ultrathin section, we used the freeze substitution and low temperature embedding procedure as described earlier (Baude *et al.*, Neuron. 1993; 11(4):771-87). Vibratome sections were cryoprotected in 1 M sucrose, frozen on a Reichert MM80E device, dehydrated in methanol at -80°C and embedded in Lowicryl HM 20 (Chemische Werke Lowi GmBH, Germany) using Leica CS auto apparatus. Ultrathin sections 80 nm thick from Lowicryl embedded blocks were picked up on nickel grids and incubated for 30 min on drops of blocking solution conisting of 1% BSA, 0.1% cold-water fish skin gelatine (Sigma), and 5% NGS in TBS containing 0.1% Triton X-100. The blocking solution was also used for diluting the primary and secondary antibodies. The grids were incubated overnight in primary antibodies (16-24 µg/ml) followed by 2 h incubation on drops of goat anti-rabbit IgG coupled to 1.4 nm gold (Nanoprobes Inc.) diluted 1:80. The antibodies were fixed with 2% glutaraldehyde for 4 min prior to silver enhancement with an HQ kit (Nanoprobes Inc.) for 3-5 min. Then sections were contrasted for electron microscopy with saturated aqueous uranyl acetate followed by lead citrate. For double-sided immunoreaction, sections were etched with sodium ethanoate for 2-3 s prior to immunoincubation (Matsubara *et al.*, Dev. Biol.1996; 180(2): 499-510).

Both preembedding and postembedding immuno-EM demonstrated unequivocally that calsyntenin-1 is located

in the postsynaptic membrane of both excitatory and inhibitory synapses. Preembedding immuno-EM with peroxidase-labeled antibodies located calsyntenin-1 in the postsynaptic membrane of synapses located on dendritic spines, dendritic shafts, and on neuronal somas (Fig. 6, A-C). In some synapses, calsyntenin-1 immunoreactivity was also found over part of the adjacent perisynaptic membranes. Rarely, floccular immunoreactivity was found in dendritic spines. Postembedding immunogold staining of rat hippocampus embedded at low temperature confirmed the localization of calsyntenin-1 in the postsynaptic membrane (Fig. 6, D-G). Both asymmetric synapses with round vesicles and thick PSDs (Type 1 according to Gray, 1959) and symmetric synapses with pleomorphic vesicles and thin PSDs (Type 2) exhibited calsyntenin-1 immunoreactivity, confirming calsyntenin-1 as a component of the postsynaptic membrane in both excitatory and inhibitory synapses.

In consideration of the postsynaptic localization of calsyntenin-1 (as shown in the present example), the calcium-binding capacity of the cytoplasmic segment of calsyntenin-1 bcomes particularly interesting. Our studies provide evidence for the presence of both high-affinity and low-affinity Ca^{2+}-binding sites. We found Ca^{2+}-binding to the cytoplasmic domain of calsyntenin-1 at a concentration as low as 0.5 μM. Therefore, the cytoplasmic domain of calsyntenin-1 binds Ca^{2+} at concentrations occurring during postsynaptic Ca^{2+} influx, suggesting calsyntenin-1 as a modulator of postsynaptic Ca^{2+} signals. In parallel to the high-capacity, low-affinity Ca^{2+}-binding function of calsequestrin, which exhibits a similar clustering of acidic residues cumulating in a contiguous stretch of 14 acidic residues, the cytoplasmic domain of calsyntenin-1 may also have the capacity for low-affinity Ca^{2+} binding. The Ca^{2+}-binding capacity of peptides with contiguous acidic residues has been linked to a general cation-binding capacity rather than specific Ca^{2+} sites. A comparison of proteins with such

acidic stretches suggested that the Ca^{2+}-binding capacity was proportional to the content of acidic residues (Lucero et al., J Biol Chem. 1994; 269(37):23112-9). X-ray crystallography suggested that the low-affinity binding of Ca^{2+} occured via intercalation of Ca^{2+} between the acidic C-terminal segments of calsequestrin dimers (Wang et al., Nat Struct Biol. 1998; 5(6):476-83). Similarly, calsyntenin-1 may bind Ca^{2+} by intercalation between its cytoplasmic moieties which are held in an ordered parallel orientation by transmembrane anchorage.

Due to its anchorage in the postsynaptic membrane, the cytoplasmic domain of calsyntenin-1 establishes a fixed Ca^{2+} buffer beneath the postsynaptic membrane. Fixed buffers, in contrast to mobile buffers, restrict the diffusion of Ca^{2+} (Kasai and Petersen, Trends Neuroscience 1994; 17(3): 95-101). They also decrease the peak values of free Ca^{2+} and, by delayed release of Ca^{2+}, prolong Ca^{2+} elevations. As a fixed Ca^{2+} buffer, calsyntenin-1 may temporarily retain Ca^{2+} in the subsynaptic zone and retard its dissipation. In this role, calsyntenin-1 may potentially be a modulatory element in synaptic processes where transient increases in intracellular Ca^{2+} are of crucial importance, such as LTP (Bliss and Collingridge, Nature. 1993; 361(6407):31-9), LTD (Linden and Connor, Annu Rev Neurosci. 1995; 18:319-57), as well as in coincidence detection within dendritic spines (Zucker, Curr Opin Neurobiol. 1999; 9(3):305-13). Postsynaptic Ca^{2+}-transients have been reported to trigger either LTP or LTD, depending on the concentration and the duration of the Ca^{2+} change. High elevations of Ca^{2+} for a few seconds induce LTP, whereas lower elevations of Ca^{2+}, lasting for a longer time span of approximately 1 min, were found to induce LTD (Malenka et al., Neuron. 1992; 9(1):121-8). The presence or absence of a Ca^{2+} buffer in the subsynaptic space could, therefore, be an important element in the mechanism determining whether the outcome of a Ca^{2+} transient is LTP or LTD. A recently reported co-

incidence detection mechanism in dendritic spines of cortical or hippocampal pyramidal neurons (Koester and Sakmann, Proc Natl Acad Sci U S A. 1998; 95(16):9596-601) generates a non-linear summation of Ca^{2+} signals, if an afferent input and a backpropagating action potential (AP) arrive at a synapse within a time window of 200 ms. When the afferent input is followed by a backpropagating dendritic AP, a supralinear summation of Ca^{2+} signals is found. In contrast, a decreased Ca^{2+} influx results when the backpropagating AP preceeds the afferent input. The enhancement of the Ca^{2+} signal that occurs when the AP follows the EPSP has been attributed to a voltage-dependent relief of the Mg^{2+} block of the NMDA receptor, whereas a Ca^{2+}-dependent NMDA-receptor inactivation has been proposed as the mechanism underlying the reduced Ca^{2+} influx when the AP arrives first at the synapse (for a review see (Zucker, Curr Opin Neurobiol. 1999; 9(3):305-13)). In both processes, a fixed buffer beneath the postsynaptic membrane could play a role. By its high-affinity Ca^{2+} binding, calsyntenin-1 might contribute to supralinear Ca^{2+} signaling. It has been suggested that buffer saturation may be an important "invisible" component in the mechanisms generating supralinear additivity of Ca^{2+} signals (Neher, Cell Calcium 1998; 24(5-6):345-57). When Ca^{2+} influx through NMDA- and voltage-gated Ca^{2+} channels coincides, more free Ca^{2+} may be generated, because the Ca^{2+} buffers are saturated by the first type of influx. By low-affinity binding of Ca^{2+} beneath the postsynaptic membrane, calsyntenin-1 could prolong Ca^{2+} transients, resulting in enhanced Ca^{2+}- dependent NMDA-receptor inactivation and, thus, a prolonged window of sublinear Ca^{2+} signaling. In a recent study with cerebellar Purkinje cells, supralinear Ca^{2+} signaling has been attributed to the saturation of a mobile high-affinity Ca^{2+} buffer (dissociation constant 0.37 μM) and to a contribution of an immobile low-affinity buffer (Maeda et al., Neuron. 1999; 24(4):989-1002). In that study, modulations of Ca^{2+} influx

or Ca^{2+} release from internal stores were excluded as the major source for the supralinearity in the Ca^{2+} responses of the Purkinje cells; rather, the supralinear responses were attributed to be predominantly due to saturation of the mobile high-affinity buffer. The immobile low-affinity buffer was suggested to contribute by prolonging the presence of Ca^{2+} and, thus, broaden the time-window of supralinear summation (Maeda et al., Neuron. 1999; 24(4):989-1002). Calsyntenin-1, by combining both high- and low-affinity Ca^{2+} buffering in one molecule at a fixed synaptic location, might contribute an important element to the coincidence detection machinery of the synapse.

Example 12:

Studies of the subcellular localization of calsyntenin-1 by subcellular fractionation and isolation of synaptosomes: Calsyntenin-1 is located in the postsynaptic membrane, but not anchored in the postsynaptic density.

To address the question whether calsyntenin-1 was firmly attached to the so-called postsynaptic density (PSD), we isolated synaptosomes by means of subcellular fractionation. For the subcellular fractionation, the protocol of Phelan and Gordon-Weeks was used (Phelan and Gordon-Weeks, 1997). Brains of 200 adult mice and 20 adult chickens, respectively, were homogenized with a Dounce homogenizer in 5 volumes of 10 mM HEPES, 0.32 M sucrose supplemented with the Mini Complete inhibitor mix (Roche). The subcellular fractionation was performed as described by Phelan and Gordon-Weeks 1997 (Isolation of synaptosomes, growth cones and their subcellular components. In: Neurochemistry - a practical approach. 2nd edition. (eds. Turner AJ, Bachelard HS) IRL Press, pp 1-38). For Western blot analysis with the antibodies R63 and R71, 100 µg total protein was loaded per lane. For controlling the correct fractionation, we used commercially available antibodies for the GluR1 subunit of the

AMPA receptor (from Pharmingen). PSD-95 is a typical component of the postsynaptic density, whereas GluR1 is a component of the AMPA-type glutamate receptor, which has been demonstrated to exhibit a firm attachment to the PSD by a high-affinity binding site on its C-ternminus (for a review: O'Brien et al., 1998). For the immunodetection of GluR1 and PSD95, 50 µg total protein were loaded per lane.

With this analysis, we found that calsyntenin-1 is a protein of the postsynaptic membrane, but does not have an intimate binding to the PSD. As demonstrated in Fig. 7, synaptosomes were enriched in full-length calsyntenin-1 and its cleavage products. Hypotonic disruption of synaptosomes and treatment with a mild detergent resulted in the solubilization of all three forms of calsyntenin-1. In contrast, typical markers of the postsynaptic density, viz. PSD-95 and GluR1 (O'Brien et al., Neuron. 1998; 21(5):1067-78), remained in the particulate fractions (P4 and PSD, according to Phelan and Gordon-Weeks 1997 (Isolation of synaptosomes, growth cones and their subcellular components. In: Neurochemistry - a practical approach. 2nd edition. (eds Turner AJ, Bachelard HS) IRL Press, pp 1-38). The clearance of calsyntenin-1 from the PSD fraction indicates that its cytoplasmic segment is not firmly associated with the subsynaptic molecular scaffold that corresponds to the PSD observed in the EM and that is operationally defined as the particulate matter resulting after detergent-treatment of synaptosomes.

Example 13:

Studies of the localization of the transmembrane cleavage product of calsyntenin-1 by immuno-electron microscopy: The transmembrane fragment of proteolytically cleaved calsyntenin-1 is accumulated in the spine apparatus of spine synapses and the subsynaptic membranes of shaft synapses

To identify the fate of the transmembrane segment of calsyntenin-1 after proteolytic cleavage, we used R71, the antibody against the membrane-proximal segment for immuno-EM with peroxidase- and gold-conjugated secondary antibodies. With peroxidase the membranes of the spine apparatus in spine synapses were labeled (Fig. 8, A and B). In some synapses, weaker immunoreactivity was also found over the postsynaptic membrane (Fig. 8A). Similarly, a strong signal was found in the lamellar membranes found beneath a fraction of the synapses in dendritic shafts and neuronal somas (not shown). With immunogold, known as less sensitive, labeling was found exclusively in association with the spine apparatus (Fig. 8, C-E) and the subsynaptic membranes of shaft synapses. Because no labeling of the subsynaptic membranous organelles was found with the N-terminal antibodies (Fig. 6), we concluded that the spine apparatus contained neither full-length calsyntenin-1 nor the N-terminal cleavage product. Therefore, the full-length as well as the 115 kD form of calsyntenin-1 found in Western blots of synaptosomes (Fig. 7) can only be derived from the postsynaptic membranes. These results indicate hat the proteolytic cleavage must occur at the cell surface, i.e. in the synaptic cleft, and that the transmembrane stump is internalized thereafter.

Proteolytic cleavage in the extracellular segment results in the release of the major extracellular portion of calsyntenin-1. This soluble fragment of calsyntenin-1 spreads in the extracellular fluids, as demonstrated by its accumulation in the cerebrospinal and the ocular vitreous fluid. The remaining transmembrane stump is internalized into the spine apparatus. Due to its membrane topology, the Ca^{2+}-binding domain of the internalized transmembrane stump covers the cytoplasmic surface of the spine apparatus. Thus, internalization may translocate the calsyntenin-1-mediated Ca^{2+} buffer from the

postsynaptic membrane to the surface of the spine apparatus.

Recently, the release of Ca^{2+} from intracellular stores, i.e. the ER and the spine apparatus, via activation of the IP3 receptors, was identified as an important contribution to the Ca^{2+} signal within dendritic spines (Finch and Augustine, Nature 1998; 396(6713):753-6). The IP3-mediated Ca^{2+} release is regulated by cytoplasmic Ca^{2+} in a biphasic mode (Taylor, Biochim Biophys Acta. 1998; 1436(1-2):19-33). Release is low at both low and high Ca^{2+} concentrations, but favored at intermediate concentrations of 200 - 300 nM. By its capacity to prolong Ca^{2+} elevations the cytoplasmic domain of calsyntenin-1 may modulate Ca^{2+} effects on IP3-mediated Ca^{2+} release.

The comparison of the results of the immuno-EM analysis and subcellular fractionation indicated that the full-length form and both cleavage products of calsyntenin-1 occur in the postsynaptic membrane (for an illustration see Fig. 9). In contrast, neither full-length calsyntenin-1 nor the N-terminal cleavage product, but exclusively the transmembrane stump, was found in internal membranes. This complete segregation of the N-terminal and the C-terminal cleavage products to the interstitial fluids and to the internal membranes, respectively, can only be explained by extracellular cleavage of calsyntenin-1 by a protease located in the synaptic cleft. The selectivity of the internalization process for the transmembrane stump of calsyntenin-1 implicates a regulatory role of the proteolytic cleavage in the synaptic cleft for the translocation of the Ca^{2+}-binding domain of calsyntenin-1 from the postsynaptic membrane to the surface of the spine apparatus.

Example 14:

Determination of the region of calsyntenin-1 bearing the proteolytic cleavage site.

The location of the proteolytic cleavage site within the sequence of full-length calsyntenin-1 remains to be determined. Based on the location of the tryptic peptides sequenced after tryptic cleavage of the released 115 kD fragment (marked in gray in Fig. 2), the released fragment must have a length of at least 747 amino acids (as counted from the N-terminus of the mature protein). Furthermore, the cleavage occurs in the extracellular moiety, i.e. on the N-terminal side of the transmembrane segment. Thus, the cleavage site has to be located after amino acid 746 (the last amino acid of the sequenced peptide number 7) and before amino acid 834 (the first amino acid of the transmembrane segment). The segment of the extracellular moiety of calsyntenin-1 defined in this way as the peptide bearing the proteolytic cleavage site is indicated below:

Cleaved sequence of chicken calsyntenin-1 (Seq. Id. No. 29):

LIRYRNWHTVS LFDRKFKLVC SELNGRYVSN EFKVEVNVIH TANPIEHANH IAAQPQFVHP VHHTFVDLSG HNLANPHPFS VVPSTATV

Cleaved sequence of human calsyntenin-1 (Seq. Id. No. 30):

LLRYRNWHAR SLLDRKFKLI CSELNGRYIS NEFKVEVNVI HTANPMEHAN HMAAQPQFVH PEHRSFVDLS GHNLANPHPF AVVPSTATV

The first amino acid is the last amino acid of the sequenced peptide number 7 of the 115 kD fragment of calsyntenin-1 (see example 2 and Fig. 2). The last amino acid is the first amino acid of the transmembrane segment (see Fig. 2).

Neither the nature of the protease that cleaves calsyntenin-1 nor the mechanism conferring selective internalization of the transmembrane stump of calsyntenin-1 are currently not known. Several extracellular proteases have been reported to be expressed in the nerv-

ous system, including tissue plasminogen activator, thrombin, neurotrypsin, and neuropsin. For two of them, namely tPA and neuropsin, transcription has been reported to be regulated by neuronal activity (Chen et al., Neurochem. Int. 1995; 26(5):455-64). They are intriguing candidates for regulators of calsyntenin-1 internalization.

Example 15:

Studies of the subcellular localization of calsyntenin-1 in growing neurons: Calsyntenin-1 is found in growth cones of growing axons during neural development

In situ hybridization indicated the expression of calsyntenin-1 mRNA in neuronal precursor cells of the germinal layers of the developing nervous system and in the early postmitotic stages of neurons in all regions of the developing nervous system. Therefore, we investigated the subcellular localization of calsyntenin-1 also in neurons during the period of cell migration and neurite growth. In this example, neurons of E18 mouse hippocampus were cultivated as described previously. When the dissociated cells from the E18 mouse hippocampus were plated at low density, their growing processes did not make contact with other cells during the first 10 days in culture. Therefore, these cultures allowed the microscopic inspection of the growth cones, the leading tips of the growing axons. As demonstrated in Fig. 10, calsyntenin-1 immunoreactivity was found on the surface of the neuronal cell soma and on both types of processes, viz. dendrites and axons. A particularly strong calsyntenin-1 immunoreactivty was found over the growth cones, as evidenced by the comparison with a double-immunofluorescence staining for calsyntenin-1 with antibody R63 (Fig. 10A) and an antibody against the axonal marker protein Tau1 (Fig. 10B). At higher magnification, calsyntenin-1 immunoreactivity in growth cones exhibited a patchy pattern,

indicating that calsyntenin-1 in growth cones occurs in multiple clusters.

The results presented in this example indicate that calsyntenin-1 is abundantly expressed on the surface of growing neurons and, thus, may have a function in developmental processes, such as neuronal migration and the formation of axons and dendrites, or in nerve regenerative functions after nervous tissue injury. The particularly strong calsyntenin-1 signal found over growth cones implicates calsyntenin-1 in growth cone functions, such as axon growth and guidance.

Example 16:

Overexpression of calsyntenin-1 in CNS neurons using transgenic mice technology

The overexpression of a gene in a transgenic mouse is a relatively direct way to study the function of a protein in vivo. For the first series of experiments chicken calsyntenin-1 was expressed under the control of the promoter of the Thy-1 gene. The Thy-1 gene is expressed in the nervous system relatively late (postnatal day 4-10, depending on the location). The expression of calsyntenin-1 under the control of the Thy-1 promoter (Gordon et al., 1987, Cell 50(3), 445-52) ensures that the earlier developmental stages are not affected. This point is essential. Calsyntenin-1 is expressed in some regions of the developing nervous system relatively early and, thus, it could play a role in early developmental functions, such as cell migration and axonal pathfinding. By using a late onset promoter it was intended to prevent perturbations of early stages of neurogenesis in the transgenic animals. However, depending on the aim of an investigation, other promoters may be used as well.

For the first transgenic mice, we chose to overpress the calsyntenin-1 of the chicken without the cytoplasmic segment, because of its potential of being selectively detected with species-specific monoclonal an-

tibodies. The chicken calsyntenin-1 exhibits an amino acid sequence identity with its counterpart of the mouse of 84.7%. Thus, a highly conserved function can be assumed.

The construct of the transgene is based on an expression vector for Thy-1 in which the translated region of Thy-1 has been substituted by a Xho-I linker (Gordon et al., 1987, Cell 50(3), 445-52). The 2682 bp long DNA fragment of chicken calsyntenin-1 used for the overexpression is derived from the chicken cDNA digested with AflIII (3 bp upstream of the start ATG) and Cac8I (9 bp downstream of the TAA stop codon). This fragment is inserted into the Thy-1 expression vector at the Xho-I linker site by a blunt-end ligation and the orientation controlled with a HindIII digest. The plasmid is rescued and the fragment to be used for the injection into the pronucleus of fertilized mouse oocytes is cut out by digestion at the two flanking Pvu1 sites. The 9 kb long injection fragment is separated on a 1% agarose gel, the band purified with a QIAEXII-kit, and the DNA eluted from the QIAEX particles with injection buffer. The generation of transgenic mice was achieved by pronuclear injection following standard protocol. The litters were screened for the presence of the transgene by PCR and Southern blotting.

By this procedure, three mouse lines overexpressing the chicken calsyntenin-1 and two lines overexpressing the mouse calsyntenin-1 were raised. The expression of the transgene was verified at the mRNA level by Northern blotting and in-situ-hybridization and at the proteine level by Western blotting. A typical overexpression was in the order of 6 to 12 fold.

By the same method, transgenic animals expressing full-length calsyntenin-1, as well as other truncated forms of calsyntenin-1 or mutated forms of calsyntenin-1 (point mutations or deletion mutations) may be generated. Instead of the Thy-1 promoter, other promoters

may be used, including promoters driving transgene expression in particular subpopulations of neurons, such as the promoter of the Purkinje cell-specific L7 protein or the limbic system-specific protease neuropsin. Alternatively, transgene expression may be put under the control of inducible promoters.

Example 17:

Expression of human calsyntenin-1 in eukaryotic (HEK293) cells

For the eukaryontic expression of human calsyntenin-1 the complete cDNA of human calsyntenin-1 was ligated into the pcDNA3.1A expression vector (Invitrogen) using the restriction sites HindIII and XbaI. The plasmid was used to transfect HEK293 cells using standard calcium phosphate transfection techniques. After 3 days the cell supernatant and cell lysate was collected, subjected to SDS-PAGE and analysed by western blotting using R63 antibody or R71 antibody. In the cell lysate of HEK293 cells the full-length human calsyntenin-1 (150 kD) was enriched whereas in the supernatant of the HEK293 cells only released human calsyntenin-1 (116 kD) was found (Fig 11).

Expression in eucaryotic cells may, alternatively, be achieved with a variety of eucaryotic expression vectors (commercially available or self-made). A frequently used eucaryotic expression system uses vectors derived from baculovirus. For eucaryotic expression, a variety of eucaryotic cell lines may be used (such as COS cells, CHO cells, HeLa cells, H9 cells, Jurkat cells, NIH3T3 cells, C127cells, CV1 cells, or Sf cells.). For a detailed description of the use of COS cells or CHO cells, or a baculovirus-based expression system see International Application Number PCT/US96/16484 or International Publication Number WO 98/16643.

Example 18:

Expression of cytoplasmic segment of calsyntenin-1 in eukaryotic (HEK293) cells

For the eukaryontic expression of calsyntenin-1 a fusion protein containing the c-kappa light chain at the C-terminus of the cytoplasmic segment of chicken calsyntenin-1 was generated. As the fusion protein was expressed as a released protein in HEK293 cells the signal peptide of NgCAM was cloned at the N-terminus of chicken calsyntenin-1. The three DNA fragments of the fusion protein, signal peptide, cytoplasmic segment and c-kappa light chain were ligated in frame using restriction sites ApaL1 and HindIII. Then the cDNA of the fusion protein was ligated into pcDNA3.1 vector (Invitrogen) using restriction sites XbaI and BamHI. The plasmid was transfected into HEK293 cells using standard calcium phosphate transfection technique. After 4 days the cell supernatant was harvested and the fusion protein was purified with a 187.1 antibody coupled affinity column.

Expression in eucaryotic cells may, alternatively, be achieved with a variety of eucaryotic expression vectors (commercially available or self-made). A frequently used eucaryotic expression system uses vectors derived from baculovirus. For eucaryotic expression, a variety of eucaryotic cell lines may be used (such as COS cells, CHO cells, HeLa cells, H9 cells, Jurkat cells, NIH3T3 cells, C127 cells, CV1 cells, or Sf cells.). For a detailed description of the use of COS cells or CHO cells, or a baculovirus-based expression system see International Application Number PCT/US96/16484 or International Publication Number WO 98/16643.

Example 19:

Cloning of the cDNA of human calsyntenin-2

The High Throughput Genomic Sequence database at NCBI was searched with the program tblastn using the protein sequence of the human calsyntenin-1. Twelve putative exons of a new member of the calsyntenin family were

found in the sequence with the accession number AC010181, which encodes a peptide homologous to the C-terminal 713 amino acids of human calsyntenin-1.

In order to determine whether these putative exons belong to a transcribed gene, PCR primers were designed based on sequence of the second putative exon (S3eBfwd: 5'-CTCCTCTGGCATCATTGACCTC-3') (Seq. Id. No. 31) and the last putative exon (S3rev: 5'-CATTTCTTCCTCGGCTTCTTCC-3') (Seq. Id. No. 32). First strand cDNA was synthesized from polyA$^+$ mRNA from human hippocampus (Clontech, catalog # 6578-1) with random hexamer primers using the ThermoScript RT-PCR System from GibcoBRL Life Technologies and used as template in a PCR reaction with the primers S3eBfwd and S3rev. A fragment of the expected length was obtained, subcloned into pBluescript KS+, and completely sequenced. Over large segments, the obtained sequence was identical with the predicted cDNA from the genomic sequence. However, an additional exon, that was not predicted from the genomic sequence, was found. This 1762 base pair fragment was radiolabeled and used as a probe to screen a human fetal brain cDNA library (Clontech 5' STRETCH PLUS in λgt10, catalog # HL3003a). 10 clones were isolated, their inserts subcloned into pBluescript KS+, and completely sequenced. The clones were assembled into a contiguous cDNA. The N-terminal 116 amino acids (by homology to calsyntenin-1) were still missing. Therefore an EcoRI-EcoNI fragment representing the most N-terminal 400 base pairs of calsyntenin-2 was used to rescreen the same cDNA library. One clone contained 109 additional N-terminal amino acids and showed a 100% identity with a sequence on another HTGS clone, AC009671. Together with this genomic clone, we were able to assemble a full-length cDNA of the human calsyntenin-2. The N-terminal sequence was confirmed with PCR using the primers hsCst2atgfwd (5'-TGCTGCGAGGATGCTGC-3' (Seq. Id. No. 33), containing the

ATG start codon) and hCs2seq4r (5'-ATGATGCCAGAGGAGGC-3') (Seq. Id. No. 34).

In summary, we found a single long ORF of 2865 nucleotides, encoding a protein of 955 amino acids. This cDNA was submitted to the EMBL/Genbank/DDBJ database under the name calsyntenin-2 and received the accession number AJ 278018. The protein translation of human calsyntenin-2 shows 57 % identity and 67 % similarity to human calsyntenin-1, and 51 % identity and 59 % similarity to human calsyntenin-3. Very much like calsyntenin-1, calsyntenin-2 is a type I transmembrane protein with a single transmembrane segment of 19 amino acids. The large N-terminal moiety of calsyntenin-2 is composed of 834 amino acids and located in the extracellular space. The C-terminal segment has a length of 102 amino acids and is highly enriched in acidic residues. Among the 102 residues of the cytoplasmic segment, 33 are acidic.

A high degree of sequence identity with calsyntenin-1 was also found in the region proximal to the transmembrane segment, which bears the protolytic cleavage site in calsyntenin-1. This suggests that calsyntein-2 also bears a proteolytic cleavage site in this segment.

The segment of calsyntenin-2 corresponding to the cleaved segment of calsyntenin-1 has the following amino acid sequence:

Putative cleaved sequence of human calsyntenin-2 (Seq. Id. No. 35):

HIRYRNWRPA SLEARRFRIK CSELNGRYTS NEFNLEVSIL HEDQVSDKEH VNHLIVQPPF LQSVHHPESR SSIQHSSVVP SIATV

In order to obtain information on the expression pattern of calsyntenin-2 in different human tissues, a Northern blot of poly(A)$^+$ RNA from adult human tissues (Cat. Nr. 7760-1, Clontech) was hybridized with a 1762 bp cDNA fragment of human calsyntenin-2 labeled with [α-^{32}P]

dCTP (Amersham) using the Prime-it II random primer labeling kit (Stratagene). Hybridization was performed for 2 h at 55 °C and the hybridization signals were analyzed with a PhosphorImager (Molecular Dynamics). A single species of calsyntenin-2 mRNA of approximately 5.5 kb was found. The highest expression of calsyntenin-2 mRNA was observed in brain, heart, and kidney. Low signals were detected in skeletal muscle. No transcript was found in placenta, lung and liver.

Example 20:

Cloning of the cDNA of human calsyntenin-3

In a database search we found an unclassified human cDNA, KIAA0726, with a sequence identity of 53.0 %, 52.3 % and 54.5 % with human, mouse, and chicken calsyntenin-1, respectively (Table I). As described in the following paragraphs, we have isolated overlapping fragments matching this sequence by RT-PCR. We found a cDNA that was identical with the sequence of KIAA0726 over a large part of the ORF, but differed in the N-terminal segment. We termed this cDNA calsyntenin-3 and submitted it to the EMBL/Genbank/DDBJ database, where it was registEred with the accession number AJ277460.

The cloning of the cDNA of human calsyntenin-3 was based on a RT-PCR strategy. In a first round, we cloned the cDNA of the mouse ortholog of KIAA0726. Approximately $2x10^6$ plaques of an adult mouse BALB/c 5' stretch plus whole brain cDNA library (ML 3000a, Clontech, Palo Alto, CA) were screened. Two independent clones with homology to KIAA0726 were isolated. A detailed analysis of these clones revealed a marked difference to KIAA0726 at the 5' end of the ORF. Both clones consist of a 120 bp 5' UTR, a translation initiation codon (ATG), and an initial part of the ORF without any homology to KIAA0726. Further downstream, however, the sequence of the clones is homologous to KIAA0726. The nucleotides adjacent to the translation codon (ATG) are in

very close agreement with the consensus sequence, as determined by Kozak (Nucleic Acids Resarch 1987; 15(20): 8125-48). Due to its homology to calsyntenin-1, the novel gene of the mouse genome was termed calsyntenin-3.

Signal peptide analysis programs predicted that mouse calsyntenin-3 contains a signal peptide of 19 aa. In contrast, with the same analysis programs, no signal peptide was predicted for KIAA0726.

A screen through the human EST database revealed a human EST (expressed sequence tag; accession number: AL133677) with a nucleotide sequence identity of 79.4 % with mouse calsyntenin-3. The 3′ region of EST AL133677 was identical with a seqment of KIAA0726 and exhibited a high degree of similarity with mouse calsyntenin-3. The 5′ region, however, exhibited a similarity with the 5′ end of the ORF of mouse calsyntenin-3, but was completely unrelated with any sequence of KIAA0726. The translated nucleotide sequence of EST AL133677 contains, like mouse Calsyntenin-3, a signal peptide of 19 aa length. The identity of the signal peptide of the ORF of EST AL133677 exhibited an amino acid sequence identity of 68.4 % with the signal peptide of mouse calsyntenin-3. Based on these characteristics, we conlcuded that the novel sequence obtained from the products of RT-PCR and EST AL133677 is the human calsyntenin-3.

In order to obtain direct information about the 5′ region of the mRNA of human calsyntenin-3, a RT-PCR approach has been undertaken. Poly $(A)^+$-selected RNA from the hippocampus of an adult human (Clontech, Palo Alto,CA) was chosen as a template for reverse transcription. First strand cDNA was obtained with the oligo (dT)-priming method (Thermoscript RT-PCR System of Life Technologies,Basel, Switzerland). *PfuTurbo* DNA Polymerase (Stratagene, La Jolla, CA) and the following primers were used to perform the PCR reaction:

forward: hSyn2UTRfor1 (starts at the 5' end of EST AL133677) (Seq. Id. No. 36)

5' CTG CAG TAG CGG GGT TG 3'

backward: RFKIA02B (ends 58 nt downstream of the TAA stop codon of KIAA0726) (Seq. Id. No. 37)

5' TGG AGT GTC TGT TTC ACC AGG 3'

This way, the complete coding sequence plus additional 275 bp of the 5' UTR of human calsyntenin-3 was obtained. DNA sequencing of both strands of the RT-PCR fragment confirmed a difference in the 5' part of human calsyntenin-3 and KIAA0726. The novel cDNA of human calsyntenin-3 contains an ORF of 2868 bp that encodes a protein of 956 amino acids consisting of a signal peptide of 19 amino acids and a transmembrane domain of 23 amino acids. The N-terminal, extracellular moiety of calsyntenin-3 is composed of 845 amino acids and the C-terminal, cytoplasmic moiety has 88 amino acids. Among the 88 amino acids of the cytoplasmic segment of calsyntenin-3, 16 have acidic side chains.

A high degree of sequence identity with calsyntenin-1 was also found in the region proximal to the transmembrane segment, which bears the protolytic cleavage site in calsyntenin-1. This suggests that calsyntein-3 also bears a proteolytic cleavage site in this segment.

The segment of calsyntenin-3 corresponding to the cleaved segment of calsyntenin-1 has the following amino acid sequence:

Putative cleaved sequence of human calsyntenin-3 (Seq. Id. No. 38):

ILRQARYRLR HGAALYTRKF RLSCSEMNGR YSSNEFIVEV NVLHSMNRVA HPSHVLSSQQ FLHRGHQPPP EMAGHSLASS HRNSMIP

The expression pattern of calsyntenin-3 mRNA in different human tissues was determined with a commercially available human multiple-tissue Northern blot (Clontech). As a probe, a 1.15 kb cDNA fragment of human calsyntenin-3, labeled with [α-^{32}P] dCTP (Amersham) using the Prime-it II random primer labeling kit (Stratagene), was used. Hybridization was performed for 2 h at 42°C and the hybridization signals were analyzed with a PhosphorImager (Molecular Dynamics). A single species of calsyntenin-3 mRNA of approximately 4 kb was revealed (Fig. 12). The highest expression of calsyntenin-3 mRNA was observed in brain. A signal of moderate intensity was found with mRNA from kidney. Low level signals were detected in pancreas, liver, heart, placenta, skeletal muscle, and lung.

The cellular resolution of the expression of calsyntenin-3 was determined by *in situ* hybridization, performed as described previously (Schaeren-Wiemers and Gerfin-Moser, Histochemistry. 1993; 100(6):431-4). On cryosections from an adult mouse brain, calsyntenin-3 mRNA was abundant in all areas of the gray matter. Inspection at higher magnification indicated a neuronal expression pattern in all areas of the CNS and the PNS. However, not all neurons expressed calsyntenin-3 mRNA and considerable differences in the expression levels were found. A very prominent example of the cell-type specific expression of calsyntenin-3 is found in the cerebellum. As shown in Fig. 12, cerebellular Purkinje cells exhibit a very strong in situ hybridization signal for calsyntenin-3, whereas all other cells of the cerebellum do not express detectable levels of calsyntenin-3 mRNA.

Example 21:

Binding of the cytosolic segment of calsyntenin-1 and the Arp2/3 complex.

While studying the scientific literature dealing with interactions between cell surface proteins and the cellular cytoskeleton, we found that the cytoplasmic part of all calsyntenin family proteins (i.e. calsyntenin-1, calsyntenin-2, and calsyntenin-3) contains at least one intriguing conserved amino acid sequence motif. This motiv consists in an acidic amino acid sequence containing a conserved tryptophan that exhibits a high degree of similarity with acidic amino acid motifs containing a conserved tryptophan in the Arp2/3 binding domain of most, if not all, of the currently known activators of the Arp2/3 complex. The acidic amino acid sequence containing a conserved tryptophan is found twice in the cytoplasmic segment of calsyntenin-1, once with the amino acid sequence ..MDWDDS.. and once with ..LEWDDS.. (amino acid sequence given in single letter code). The cytoplasmic sequence of calsyntenin-2 contains one ..MDWDDS.. and one ..LEWDDS.., and the cytoplasmic segment of calsyntenin-3 contains a single motif of this kind, namely ..LFWDDS... The Arp2/3 complex plays a central role in the regulation of the actin-based cellular motility, by regulating actin filament growth and branching (for reviews see: Borisy and Svitkina, Curr. Opin. Cell Biol. 12: 104-112, 2000; Pantaloni et al., Science 292: 1502-1506; Higgs and Pollard, Annu. Rev. Biochem., 70: 649-676, 2001; and references therein). Arp2/3 activators containing a similar acidic motif with a conserved tryptophan include human WASP (Abbreviation for: Wiscott Aldrich Syndrome Protein), the related human N-WASP, the human Scar/WAVE1 proteins, and cortactin, exhibiting the sequences ..DDEWDD, ..DDEWED and ..EVDWLE, and ..ADDWET.., respectively (for WASP, N-WASP, and Scar/WAVE1 see Higgs and Pollard, Annu. Rev. Biochem. 70: 649-676, 2001; for cortactin see Uruno et al., Nature Cell Biol. 3: 259-266, 2001). The importance of the conserved tryptophan and the adjacent acidic amino acids for Arp2/3 binding and the Arp2/3 function in actin poyme-

rization has been demonstrated by site-directed mutagenesis of cortactin (Uruno et al., Nature Cell Biol. 3: 259-266, 2001). Site directed mutagenis of both the tryptophan and the two amino acid residues preceding the tryptophan in the sequence ..ADDWET.. resulted in the loss of Arp2/3 binding and Arp2/3-mediated actin polymerization. All these Arp2/3 activator proteins are resident in the cytoplasm and have been reported to link intracellular signals generated by the transmembrane signaling of receptors for extracellular regulators, such as growth factor, cytokines, etc., into activation of the Arp2/3 complex. A crucial intermediate step in the signaling cascade from activated transmembrane receptors to the activation of the Arp2/3 activators has been attributed to the small GTP-binding proteins of the Rho family (for a review: Takai et al., Physiol. Rev. 81:153-207, 2001). Activated Arp2/3 complex in turn initiates the generation of new actin filaments and the branching of pre-existing actin filaments (for reviews see: Borisy and Svitkina, Curr. Opin. Cell Biol. 12: 104-112, 2000; Pantaloni et al., Science 292: 1502-1506; Higgs and Pollard, Annu. Rev. Biochem., 70: 649-676, 2001; and references therein). As a result of the enhanced cytoskeletal dynamics, the cells generate and/or retract plasma membrane protrusions, such as filopodia and lamellipodia (Borisy and Svitkina, Curr. Opin. Cell Biol. 12: 104-112, 2000). In the growing tip of the axons growing out of neurons, termed growth cones, the enhanced activity so generated translates into an enhanced exploratory activity and enhanced axon growth and pathfinding activity (Hu and Reichardt, Neuron 22, 419-422, 1999; Suter and Forscher, Curr. Opin. Neurobiol. 8: 106-116, 1998; Dickson, Curr. Opin. Neurobiol. 11: 103-110, 2001). The enhanced dynamics of actin filaments in the dendritic spines of neurons of the central nervous system results in an enhanced motility, which in turn may regulate the morphological shape and the electrical properties of the spine. As a

consequence, the postsynaptic response to presynaptic signals may be altered (Segal et al., Trends Neurosci. 23: 53-57, 2000; Halpain, Trends Neurosci. 23: 141-146, 2000; Matus, Science 290: 754-758, 2000; Scott and Luo, Nature Neurosci. 4: 359-365, 2001). In non-neuronal cells, the enhanced dynamics of actin filaments induced via Arp2/3 activation results in an increase in cell motility, accompanied by a boost in the formation of membrane protrusions, such as lamellipodia, and enhanced migratory activity (Holt and Koffer, Trends Cell Biol. 11: 38-47, 2001; Mullins, Curr. Opin. Cell Biol. 12: 91-96, 2000; Prokopenko et al., J. Cell Biol. 148: 843-848, 2000). A dysregulated signalling from the cell surface to the cytoskeleton, resulting in altered cell motility, enhanced formation of lamellipodia, and enhanced locomotion, when found in tumour cells, strengthens the capacity of the tumor cells for invasive growth and metastasis (Radisky et al., Seminars Cancer Biol. 11:87-95, 2001; Kassis et al., Seminars Cancer Biol. 11:105-119, 2001; Condeelis et al., Seminars Cancer Biol. 11:119-128, 2001; Price and Collard, Seminars Cancer Biol. 11:167-173, 2001).

The acidic sequence containing a tryptophan residue was also found to be crucial for the induction and the branching of actin filaments generated Listeria monocytogenes (Higgs and Pollard, Annu. Rev. Biochem., 70: 649-676, 2001; Cameron et al., Curr. Biol. 11: 130-135, 2001). After invading the cytosol of the host cell, these bacteria use the cellular actin machinery for their own locomotion. The bacterial surface protein ActA initiates the formation of actin filaments on the surface of Listeria monocytogenes. Very much like the cells own Arp2/3 activators, ActA of Listeria monocytogenes contains a tryptophan flanked by acidic residue. It has the amino acid sequence ..DEWEE.. (for a review: Higgs and Pollard, Annu. Rev. Biochem. 70: 649-676, 2001).

Based on the high degree of similarity with the Arp2/3-binding and Arp2/3-activating motif of the

currently published Arp2/3 activators, we speculated that the proteins of the calsyntenin family may regulate the dynamics of the actin cytoskeleton via binding to and regulating of the function of the Arp2/3 complex. To test this hypothesis, we generated a fusion protein composed of glutathion-S-transferase (GST) and the cytoplasmic segment of calsyntenin-1 (GST-Cst_C). The COOH-terminal domain (Cst_C) of human calsyntenin-1 was expressed and purified as a GST fusion protein in E. coli. The region encoding the cytoplasmic segment of human calsyntenin-1 (Cst_C: residues 881-981) was amplified by PCR from human brain cDNA using oligonucleotides hsCst1-881f (5'-CGGGATCCCGCATCCGGGCCGCACAT-3') (Seq. Id. No. 39) and hsCst1-981r (5'-GGGAATTCCTCAGTAGCTGAGGGTGGAG-3') (Seq. Id. No. 40) as forward and reverse primer, respectively. The Cst_C PCR fragment was cloned into the BamH1/EcoR1 sites of the pGEX6P-1 plasmid. The resulting pGEX-GST-Cst_C plasmid was transfected into the E. coli strain BL21 and the expression of GST-Cst_C protein was induced according to standard procedures. GST-Cst_C fusion protein was purified according to standard procedures using Glutathione-Sepharose (Amersham Pharmacia Biotech) and kept at 4°C until use. To generate an affinity column, 3.5 mg of GST-Cst_C were bound to 0.75 ml Glutathion-Sepharose by batch incubation. Thereafter, the GST-Cst_C conjugated Glutathion-Sepharose was packed into a column and equilibrated with Buffer B (see below). Bovine brain extract was prepared according to the procedure described previously (Uruno et al., Nature Cell Biol. 3:259-266, 2001). Briefly, 100 g of frozen bovine brain were minced with a Waring blender in 100 ml of buffer Q (20 mM Tris, 100 mM NaCl, 5 mM MgCl2, 5 mM EGTA and 1 mM dithiothreitol (DTT), pH 8.0), supplemented with 50 µg/ml phenylmethylsulphonyl fluoride, 5 µg/ml leupeptin and 1 µg/ml aprotinin. The minced tissue was further homogenized using a Dounce homogenizer and was clarified by centrifugation at 10,000g for 60 min at 4 °C. The supernatant was

subjected to chromatography in a 100-ml Q Sepharose Fast-flow column equilibrated with buffer Q. The flow-through was collected, supplemented with 0.1 mM ATP, and loaded on a GST-Cst_C Glutathione-Sepharose column equilibrated with buffer B (50 mM Tris, 25 mM KCl, 1 mM MgCl2, 0.5 mM EDTA, 1 mM DTT, 0.1 mM ATP, pH 7.5). After washing with buffer B, elution was initiated with 0.2 M KCl in buffer B, followed by a second elution with buffer B containing 0.2 M MgCl2. The eluted protein was analyzed by SDS-PAGE followed by Western blotting using a commercially available antibody against the Arp3 subunit of the Arp2/3 complex (Santa Cruz). As demonstrated in Figure 13, the Arp2/3 complex, contained in the complex mixture of proteins extracted from bovine brain, bound to the immobilized GST-Cst_C fusion protein and was released only when elution conditions were applied (Elution A in Figure 13). No binding of Arp2/3 was observed when bovine brain extract was passed over a column containing only the GST part. This indicates that the observed binding of Arp2/3 to the GST-Cst_C fusion protein is mediated by the cytoplasmic segment of calsyntenin-1. Based on the presence of highly similar segments containing a conserved tryptophan flanked by acidic amino acids in the cytoplasmic parts of calsyntenin-2 and calsyntenin-3 all members of the calsyntenin family bind to and, thus, regulate Arp2/3 activity.

To demonstrate direct binding between the cytoplasmic segment of calsyntenin-1 and the Arp2/3 complex, Arp2/3 complex was purified according to a published protocol (Egile et al., J. Cell Biol. 146:1319-1332, 1999). To prepare an affinity ligand for Asp2/3 complex, the COOH-terminal domain (VCA) of human N-WASP was expressed as a GST fusion protein in E. coli and purified using a Glutathion-Sepharose column. The region encoding the VCA segment of human N-WASP (residues 392-505) was amplified by PCR from human brain cDNA using oligonucleotides phNW392 (5'-ccggaattcCCTTCTGATGGGGAC

CATCAG-3') (Seq. Id. No. 41) and phNW505 (5'-ccgctcgag TCAGTCTTCCCACTCAT CATC-3') (Seq. Id. No. 42) as forward and reverse primer, respectively, as described previously (Egile et al., J. Cell Biol. 146:1319-1332, 1999). The PCR fragment encoding N-WASP VCA was cloned into the XhoI site of the pGEX6P-1 plasmid, to generate the pGEX-VCA plasmid. GST-VCA protein was expressed in the E. coli strain BL21 according to standard induction and purification procedures. GST-VCA fusion protein was purified on a Glutathion-Sepharose column and eluted following the protocol recommended by the supplier (Amersham Pharmacia Biotech). To generate an affinity column, purified GST-VCA was bound to Glutathion-Sepharose beads by batch incubation. GST-VCA glutathione Sepharose beads were stored at 4°C until use.

The GST-VCA Glutathion-Sepharose was used as the affinity matrix to purify the Arp2/3 complex, as described previously (Uruno et al., Nature Cell Biol. 3:259-266, 2001). Briefly, 100 g of frozen bovine brain were minced with a Waring blender in 100 ml buffer Q (20 mM Tris, 100 mM NaCl, 5 mM MgCl2, 5 mM EGTA, 1 mM dithiothreitol, pH 8.0) supplemented with 50 µg/ml phenylmethylsulphonyl fluoride, 5 µg/ml leupeptin and 1 µg/ml aprotinin. The minced tissue was further homogenized using a Dounce homogenizer and was clarified by centrifugation at 10,000g for 60 min at 4 °C. The supernatant was subjected to chromatography in a 100-ml Q Sepharose Fast-flow column equilibrated with buffer Q. The flow-through containing the Arp2/3 complex was collected, supplemented with 0.1 mM ATP and fractionated on a GST-VCA glutathione-sepharose column equilibrated with buffer B (50 mM Tris, 25 mM KCl, 1 mM MgCl2, 0.5 mM EDTA, 1 mM DTT, 0.1 mM ATP, pH 7.5). After washing with 0.2 KCl in buffer B, the Arp2/3 complex was eluted with buffer B containing 0.2 M MgCl2. The protein was then dialysed against buffer B and concentrated with a Centriprep 10 cartridge. The concentrated Arp2/3 complex was stored in buffer B con-

taining 30% glycerol at -80 °C. Protein concentration was determined by the BCA method (Pierce protein assay), using BSA as a standard.

To investigate whether the cytoplasmic segment of calsyntenin-1 has the capacity to bind the Arp2/3 complex, the GST-Cst_C fusion protein was bound to Glutathion-Sepharose beads by batch incubation. For a control, the GST-VCA fusion protein, which is an established ligand of the Arp2/3 complex (Uruno et al., Nature Cell Biol. 3:259-266, 2001) and GST alone were bound to Glutathion-Sepharose. GST-Cst_C, GST-VCA, or GST (5 µg), immobilized on Glutation-Sepharose beads, were mixed with 10 pmol of purified Arp2/3 complex in buffer A (100 µl of 50 mM Tris, 1 % Triton-X-100), and incubated for 2 h at 4 °C on a rotating wheel. The beads were rinsed three times with buffer A and then boiled in two times SDS sample buffer. The resulting sample buffer was loaded on an SDS-PAGE gel. The electrophoretically separated proteins were electrotransferred onto nitrocellulose using standard protocols. The Arp2/3 complex was visualized with a polyclonal anti-Arp3 antibody (purchased from Santa Cruz) according to standard immunoblotting procedures. We found unequivocal evidence for direct binding of Arp2/3 to GST-Cst_C and GST-VCA, but not GST alone. These results demonstrate a direct binding interaction between the cytoplasmic segment of calsyntenin-1 and the Arp2/3 complex.

In summary, calsyntenin-1 containing the cytoplasmic sequences ..MDWDDS.. and ..LEWDDS.. is capable of binding the Arp2/3 complex. This indicates that calsyntenin-1 uses the same binding site to interact with the Arp2/3 complex at as the currently known regulators of Arp2/3 activity, including human WASP, human N-WASP, the human Scar/WAVE1 proteins, cortactin, and the ActA protein of Listeria monocytogenes, in which the binding site comprising the conserved tryptophan includes the sequences ..DEWDD, ..DEWED, ..VDWLE, ..ADDWET.., and

..DEWEE, respectively (for an overview see: Higgs and Pollard, Annu. Rev. Biochem. 70: 649-676, 2001 and references therein). Thus, calsyntenin-1, by means of its cytoplasmic part competes with established regulators of Arp2/3 activity and, in doing so, takes part in the regulation of Arp2/3 activity. Calsyntenin-1 is the first regulator of the Arp2/3 complex that is a transmembrane protein. In contrast, the currently known Arp2/3 regulators, including WASP, N-WASP, the proteins of the Scar/WAVE1 family, and cortactin are cytoplasmic proteins and depend on other intracellular mediators of extracellular signals. Calsyntenin-1 may transduce extracellular signals received by ist extracellular part directly into an activity-regulating signals to the Arp2/3 complex. The presence of highly similar conserved acidic segments containing a conserved tryptophan in the cytoplasmic parts of calsyntenin-2 and calsyntenin-3 indicates that all members of the calsyntenin family members bind to and, thus, regulate Arp2/3 activity.

While there are shown and described presently preferred embodiments of the invention, it is to be distinctly understood that the invention is not limited thereto but may be otherwise variously embodied and practiced within the scope of the following claims.

PATENT CLAIMS

1. An isolated polypeptide for the use as a pharmaceutical comprising an amino acid sequence at least 50% identical to a sequence selected from the group consisting of:

a) a full length amino acid sequence selected from the group consisting of Seq. Id. No. 2, Seq. Id. No. 4 and Seq. Id. No. 6,

b) a polypeptide comprising at least one, preferably two, most preferably three of amino acid sequences selected from the group consisting of: sequence of residues from about 46 to about 165, sequence of residues 166 to about 257, sequence of residues from about 774 to about 861 and sequence of residues from about 881 to about 981 of Seq. Id. No. 2,

c) a polypeptide comprising at least one, preferably two, most preferably three of amino acid sequences selected from the group consisting of: sequence of residues from about 66 to about 158, sequence of residues from about 182 to about 259, sequence of residues from about 751 to about 834 and sequence of residues from about 854 to about 955 of Seq. Id. No. 4,

d) a polypeptide comprising at least one, preferably two, most preferably three of amino acid sequences selected from the group consisting of: sequence of residues from about 51 to about 142, sequence of residues from about 167 to about 244, sequence of residues from about 759 to about 845 and sequence of residues from about 869 to about 956 of Seq Id. No. 6

e) a polypeptide comprising at least one, preferably two, most preferably three of amino acid sequences selected from the group consisting of: sequence of residues from about 881 to about 981 of Seq. Id. No. 2, sequence of residues from about 854 to about 955 of Seq. Id. No. 4 and sequence of residues from about 869 to about 956 of Seq Id. No. 6,

and having calcium binding activity and/or the capacity to bind to the Arp2/3 complex.

2. The polypeptide of claim 1 wherein the polypeptide sequence is at least 60% identical and more preferably more than 65% identical to an amino acid sequence selected from the group consisting of:

a) a full length amino acid sequence selected from the group consisting of Seq. Id. No. 2, Seq. Id. No. 4 and Seq. Id. No. 6

b) a polypeptide comprising at least one, preferably two, most preferably three of amino acid sequences selected from the group consisting of: sequence of residues from about 46 to about 165, sequence of residues 166 to about 257, sequence of residues from about 774 to about 861 and sequence of residues from about 881 to about 981 of Seq. Id. No. 2,

c) a polypeptide comprising at least one, preferably two, most preferably three of amino acid sequences selected from the group consisting of: sequence of residues from about 66 to about 158, sequence of residues from about 182 to about 259, sequence of residues from about 751 to about 834 and sequence of residues from about 854 to about 955 of Seq. Id. No. 4,

d) a polypeptide comprising at least one, preferably two, most preferably three of amino acid sequences selected from the group consisting of: sequence of residues from about 51 to about 142, sequence of residues from about 167 to about 244, sequence of residues from about 759 to about 845 and sequence of residues from about 869 to about 956 of Seq Id. No. 6,

e) a polypeptide comprising at least one, preferably two, most preferably three of amino acid sequences selected from the group consisting of: sequence of residues from about 881 to about 981 of Seq. Id. No. 2, sequence of residues from about 854 to about 955 of Seq. Id. No. 4 and sequence of residues from about 869 to about 956 of Seq Id. No. 6,

and having calcium binding activity and/or the capacity to bind to the Arp2/3 complex.

3. An isolated polypeptide for the use as a pharmaceutical comprising an amino acid sequence selected from

sequences comprising a stretch of at least 100 amino acids with a minimal identity percentage of 50%, preferably 55% and more preferably 60% to a amino acid sequence selected from the group consisting of Seq. Id. No. 2, Seq. Id. No. 4 and Seq. Id. No. 6, said sequences having calcium binding activity and/or the capacity to bind to the Arp2/3 complex.

4. The polypeptide according to anyone of the preceding claims which

is a transmembrane protein and which

is expressed predominatly in cells of the nervous system.

5. The polypeptide according to claim 4 which is expressed in neurons.

6. The polypeptide according to claim 5 which is localised to the postsynaptic membrane of synapses.

7. The polypeptide according to claim 6 which is localized in a membrane of a spine apparatus of spine synapses and in a membrane of subsynaptic endoplasmatic reticulum of shaft synapses.

8. The polypeptide according to anyone of claims 4 to 7 which is expressed in tumor cells.

9. The polypeptide according to anyone of the preceding claims which has its major calcium-binding domain in the cytoplasmic compartment.

10. The polypeptide according to anyone of the preceding claims which has at least one binding site for the Arp2/3 complex.

11. An isolated nucleotide sequence encoding a polypeptide according to anyone of the preceding claims for the use as pharmaceutical.

12. An isolated nucleotide sequence encoding a polypeptide as defined in anyone of the preceding claims which has, due to at least one point mutation, insertion or deletion, lost its function.

13. The nucleotide sequence according to claim 12 for the use as a diagnostic tool.

14. A pharmaceutical composition comprising a polypeptide according to anyone of claims 1 to 10.

15. Pharmaceutical composition comprising a polypeptide as defined in anyone of claims 1 to 10 and/or a nucleotide sequence according to claim 11.

16. Use of a polypeptide or a partial sequence thereof as defined in anyone of the preceding claims for the use as a tool for the development of a pharmaceutical.

17. Use of a protein or a DNA sequence as defined in anyone of the preceding claims for the screening and for the preparation of a medicament for the treatment of disorders, in particular disorders of the nervous system, more particular of the central nervous system, most preferably the brain.

18. Use according to claim 17, characterized in that said disorders, in particular of the nervous system, more particular of the brain, are disorders due to lack of cleavage or miscleavage or excessive cleavage of a protein as defined in one of claims 1-11 induced by at least one protease.

19. Use according to claim 18, characterized in that the protease is tissue-type plasminogen activator, abbreviated as tPA, urokinase-type plasminogen activator, abbreviated as uPA, or plasmin or neurotrypsin, or thrombin, or neuropsin.

20. Use according to anyone of claims 17 to 19, characterized in that said disorders, in particular of the nervous system, are due to perturbed processing of intracellular calcium signals.

21. Use according to claim anyone of claims 17 to 20, characterized in that said disorders, in particular of the nervous system, are due to perturbed processing of extracellular signals that regulate the cellular motility processes by means of regulating the activity of the Arp2/3 complex.

22. Use according to one of claims 17 to 21, characterized in that the medicament is a medicament for the minimization of the tissue destruction during and/or after stroke.

23. Use according to anyone of claims 17 to 22, characterized in that the medicament prevents the cell death of cells of the nervous system.

24. Use of a DNA sequence or a protein as defined in anyone of claims 1-10 for the preparation of a medicament for the treatment of tumors, including prevention or reduction of the growth, the expansion, the infiltration and the metastasis of primary and metastatic tumors, in particular brain tumors or tumors of the retina.

25. Use according to claim 24, characterized in that said tumors involve in their growth, expansion, infiltration, metastasis and promotion of blood vessels or neoangiogenesis an enhanced activity of the Arp2/3 complex.

26. Use according to claim 25, characterized in that said enhanced activity of the Arp2/3 complex is mediated by an abnormal or excessive or reduced regulatory function of one of the sequences as defined in anyone of claima 1-10.

27. Use according to anyone of claims 24 to 26, characterized in that said tumors involve in their growth, expansion, infiltration, metastasis and promotion of blood vessels or neoangiogenesis at least one protease functionally connected with a polypeptide as defined in anyone of the claims 1-10.

28. Use according to claim 27, characterized in that the protease is a member of one of the following protease families:

- Serine Protease family such as tissue-type plasminogen activator (tPA), urokinase-type plasminogen activator (uPA), plasmin, thrombin, neurotrypsin, neuropsin, elastases, cathepsin G,

- Matrix Metalloproteinases family such as collagenases, gelatinases, stromelysins, matrylisins,

- Cystein Proteases family such as cathepsin B and cathepsin D.

29. A method for the production of polypeptides as defined in anyone of claims 1 to 10 or such polypeptide expressing cells, characterized in that suitable host cells are transfected with a DNA sequence as defined in claim 12 in a vector ensuring the expression of said DNA sequence in said host cell, and in that said transfected cells are cultured under suitable conditions allowing said expression.

30. A synthetic or chemical method for the production of polypeptides, peptides or nucleic acid sequences representing at least part of the sequences defined in claims 1 to 13 and having the ability to mimic or to block, respectively, the biological activity of calsyntenin, in particular the calcium binding activity.

31. Use of the DNA sequences and/or the polypeptides as defined in anyone of claims 1 to 13 as tools in the screening of pharmaceutical drugs.

32. Use of a sequence as defined in claim 11 as a means to produce antigens or as antigen for the production of antibodies.

33. Transgenic non human animal, characterized in that it comprises an exogenous DNA sequence as defined in claim 12 in an environment allowing protein expression.

34. Use of a DNA sequence as defined in claim 12 or 13 or fragments thereof for the preparation of a

diagnostic preparation for the diagnosis of disorders due to defects in the genomic sequence comprising a DNA sequence according to claim 11.

35. A vector or artificial chromosome comprising a DNA sequence as defined in claim 12 for the use in gene therapeutical applications in humans and in animals.

36. An isolated polypeptide comprising an amino acid sequence which is at least 60% identical to the amino acid sequence of Seq.Id. No. 4.

37. The polypeptide of claim 36 wherein the amino acid sequence is identical to Seq. Id. No. 4.

38. An isolated polypeptide comprising an amino acid sequence which is at least 98.5 % identical to the amino acid sequence of Seq. Id. No. 6.

39 The polypeptide of claim 37 which is identical to Seq. Id. No. 6.

40. A nucleotide sequence encoding a polypeptide according to anyone of claims 36 to 39.

41. A protease which cleaves a polypeptide as defined in anyone of claims 1 to 10 in its extracellular part.

42. A cell extract comprising a protease which cleaves a polypeptide as defined in anyone of claims 1 to 10.

Figure 1

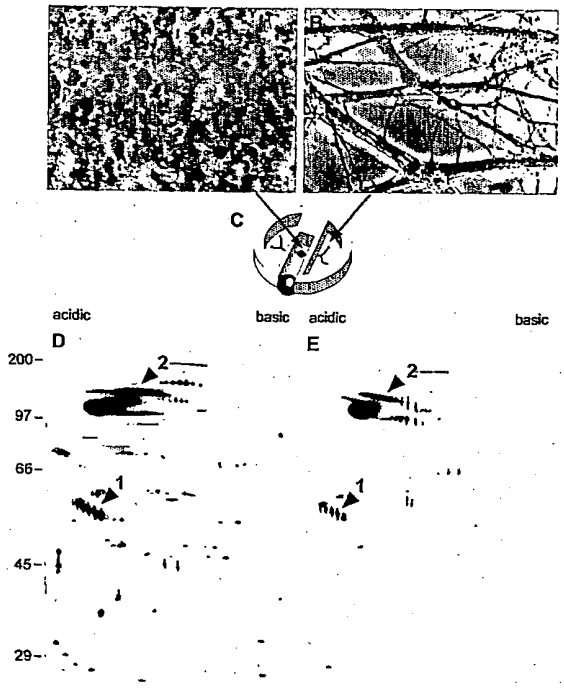


Figure 2

Figure 3

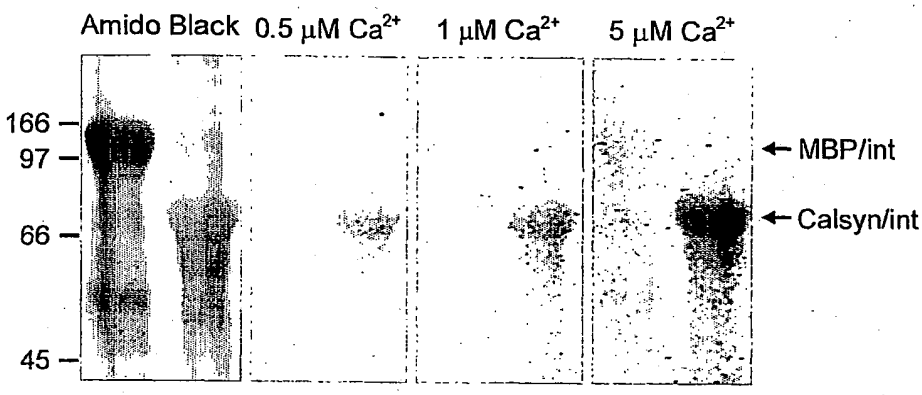

Amido Black
0.5 μM Ca2+
1 μM Ca2+
5 μM Ca2+
166
97
66
45
MBP/int
Calsyn/int

Figure 4

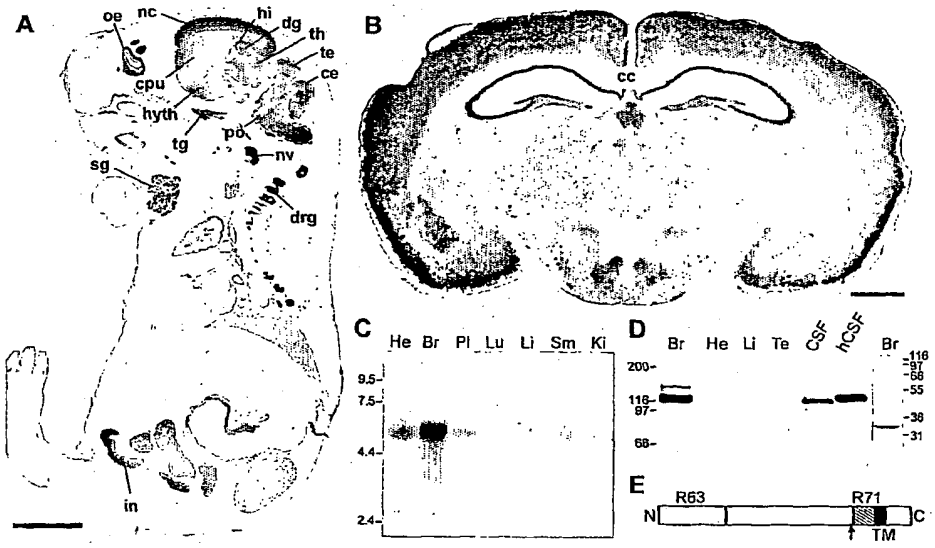

A
oe
nc
hi
dg
th
te
ce
cpu
hyth
tg
po
nv
sg
drg
in
B
cc
C
He Br Pl Lu Li Sm Ki
9.5
7.5
4.4
2.4
D
Br He Li Te CSF hCSF Br
200
116
97
68
55
38
31
E
R63
R71
N
C
TM

Figure 5

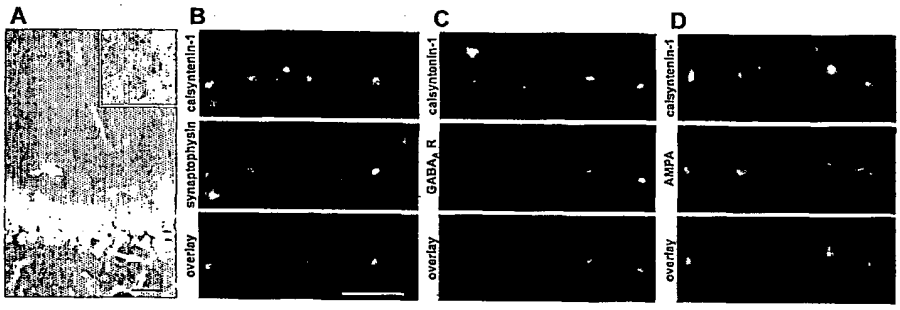

A
B
C
D
calsyntenin-1
synaptophysin
overlay
calsyntonin-1
GABA, R
overlay
calsyntenin-1
AMPA
overlay

Figure 6

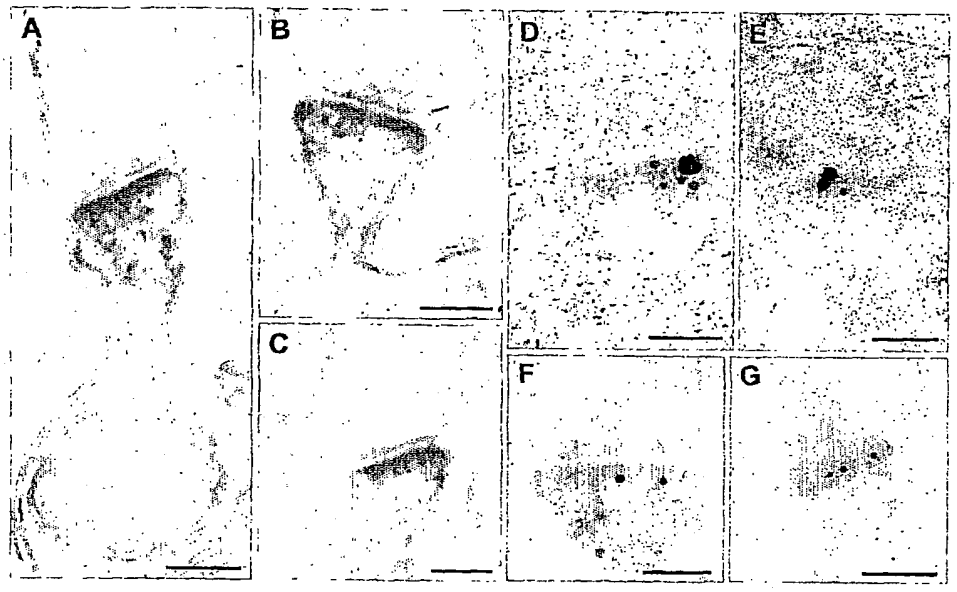


Figure 7

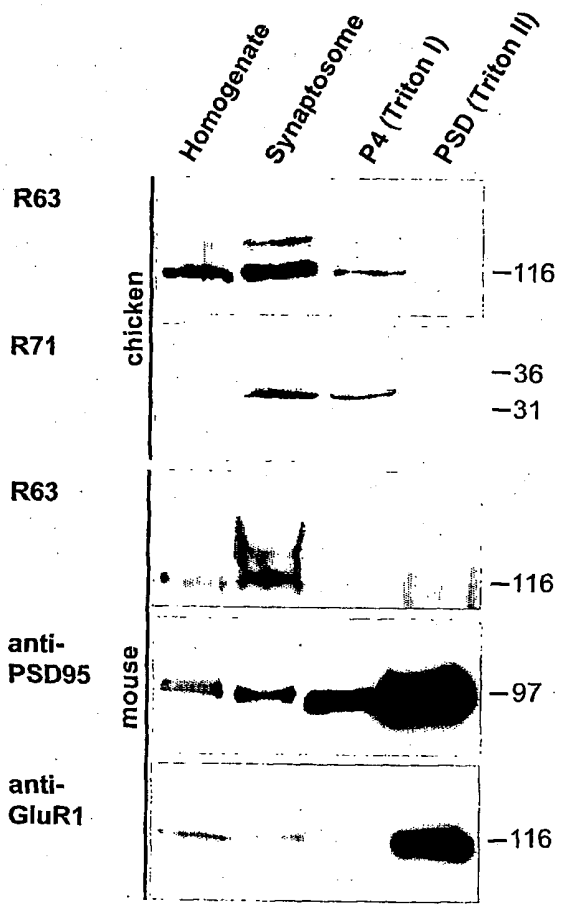

Homogenate
Synaptosome
P4 (Triton I)
PSD (Triton II)
R63
chicken
R71
R63
anti-PSD95
mouse
anti-GluR1
116
36
31
116
97
116

Figure 8

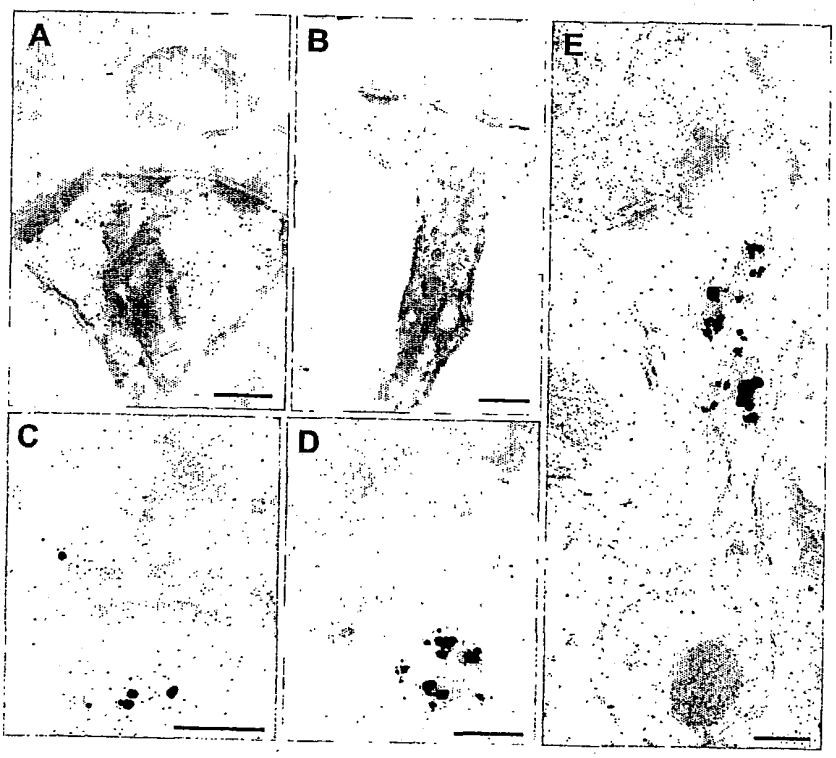


Figure 9

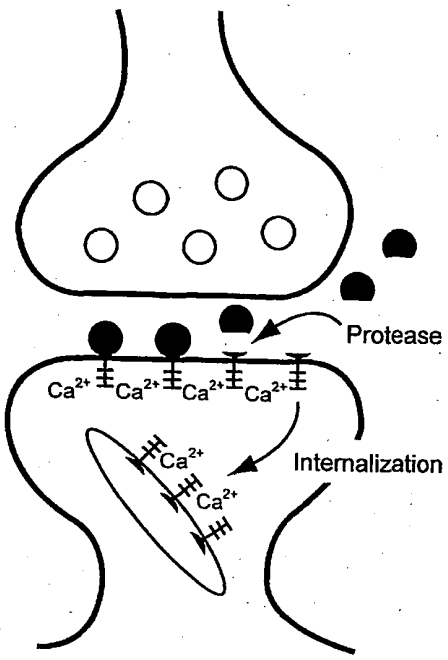


Figure 10

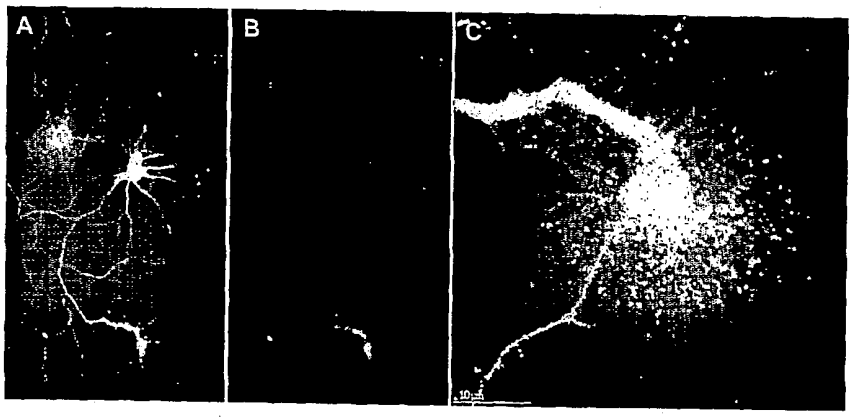


Figure 11

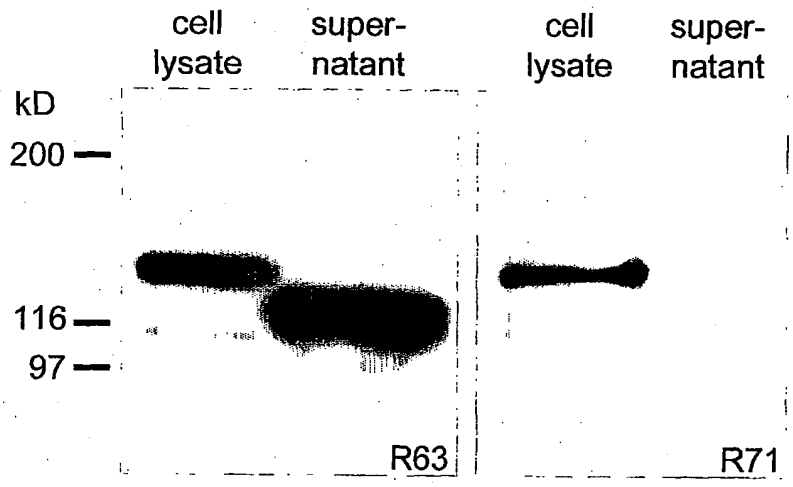

cell lysate
super-natant
cell lysate
super-natant
kD
200
116
97
R63
R71

Figure 12

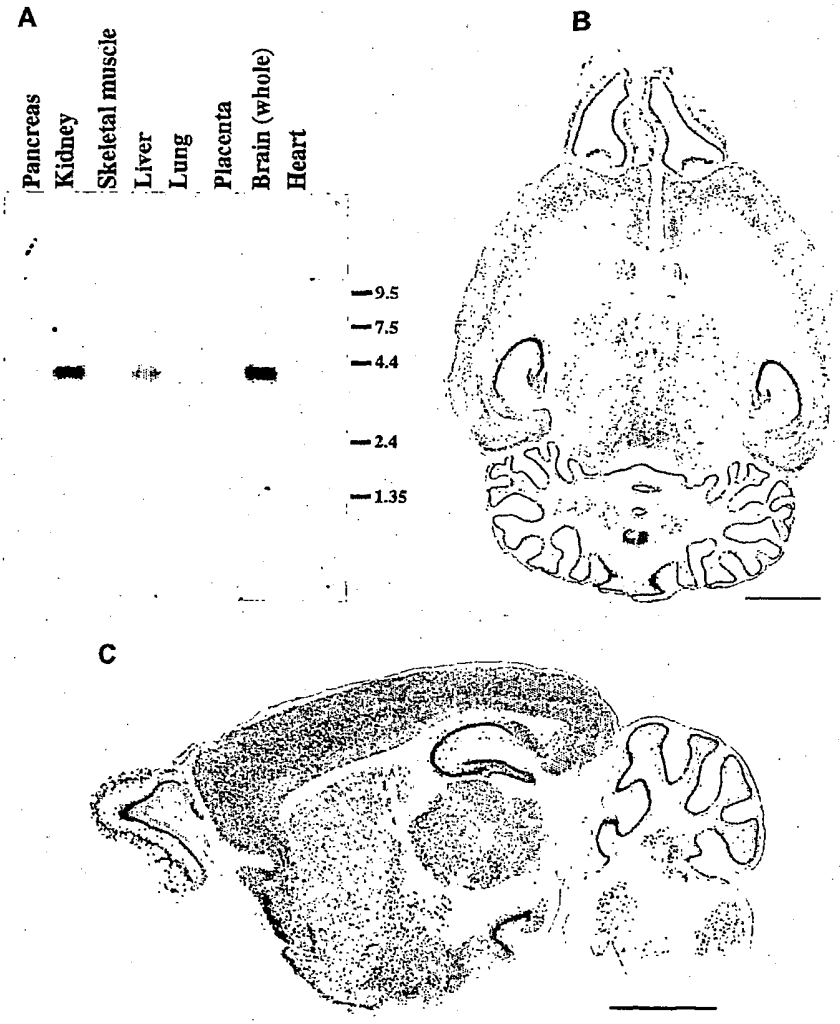
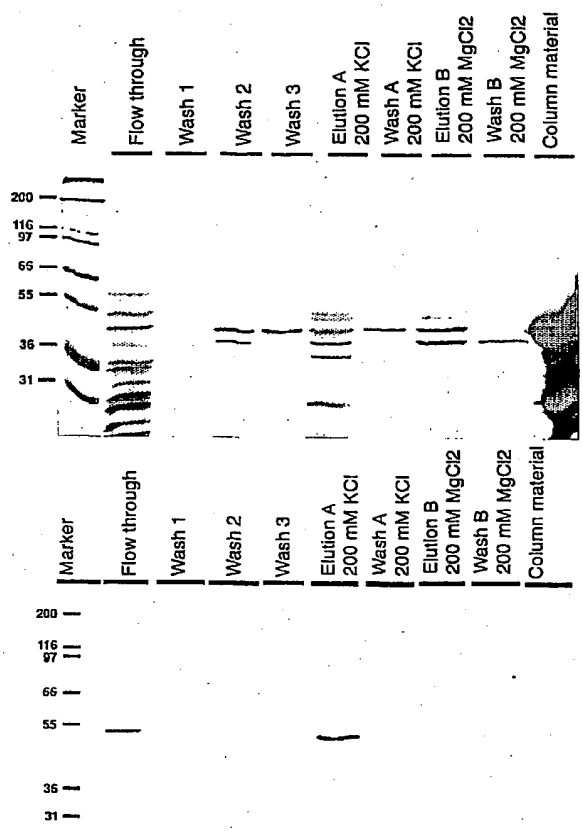

A
Pancreas
Kidney
Skeletal muscle
Liver
Lung
Placenta
Brain (whole)
Heart
9.5
7.5
4.4
2.4
1.35
B
C

Figure 13


Marker
Flow through
Wash 1
Wash 2
Wash 3
Elution A 200 mM KCl
Wash A 200 mM KCl
Elution B 200 mM MgCl2
Wash B 200 mM MgCl2
Column material
200
116
97
66
55
36
31

SEQUENCE LISTING

<110> Universität Zürich

<120> Calcium binding proteins

<130> Calsyntenins

<140>
<141>

<150> EP 00810830.0
<151> 2000-09-14

<160> 42

<170> PatentIn Ver. 2.1

<210> 1
<211> 3700
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (236)..(3178)

<400> 1
gaattccgga gttgctgccg ctgccttcag caagacgctg ctctgaggcg gggagggcgc 60

cgcgtcctga gcgcgcggcc cagcgtcacg gcggcggcgg cggcggctcc tccttggacc 120

cccggagctc cccgcgccgc gagcagctgg ccccaggccc ctagagcccc gagagctccg 180

agagctccgc tcggcgtccc gcgcgcctcc ctgccgctcc cgccccgggc tggcg atg 238
Met
1

ctg cgc cgc ccc gct ccc gag ctg gcc ccg gcc gcc cgg ctg ctg ctg 286
Leu Arg Arg Pro Ala Pro Glu Leu Ala Pro Ala Ala Arg Leu Leu Leu
5 10 15

gcc ggg ctg ctg tgc ggc ggc ggg gtc tgg gcc gcg cga gtt aac aag 334
Ala Gly Leu Leu Cys Gly Gly Gly Val Trp Ala Ala Arg Val Asn Lys
20 25 30

cac aag ccc tgg ctg gag ccc acc tac cac ggc ata gtc aca gag aac 382
His Lys Pro Trp Leu Glu Pro Thr Tyr His Gly Ile Val Thr Glu Asn

```
    35                  40                  45

gac aac acc gtg ctc ctc gac ccc cca ctg atc gcg ctg gat aaa gat   430
Asp Asn Thr Val Leu Leu Asp Pro Pro Leu Ile Ala Leu Asp Lys Asp
 50                  55                  60                  65

gcg cct ctg cga ttt gca gag agt ttt gag gtg aca gtc acc aaa gaa   478
Ala Pro Leu Arg Phe Ala Glu Ser Phe Glu Val Thr Val Thr Lys Glu
                 70                  75                  80

ggt gag att tgt gga ttt aaa att cac ggg cag aat gtc ccc ttt gat   526
Gly Glu Ile Cys Gly Phe Lys Ile His Gly Gln Asn Val Pro Phe Asp
             85                  90                  95

gca gtg gta gtg gat aaa tcc act ggt gag gga gtc att cgc tcc aaa   574
Ala Val Val Val Asp Lys Ser Thr Gly Glu Gly Val Ile Arg Ser Lys
        100                 105                 110

gag aaa ctg gac tgt gag ctg cag aaa gac tat tca ttc acc atc cag   622
Glu Lys Leu Asp Cys Glu Leu Gln Lys Asp Tyr Ser Phe Thr Ile Gln
    115                 120                 125

gcc tat gat tgt ggg aag gga cct gat ggc acc aac gtg aaa aag tct   670
Ala Tyr Asp Cys Gly Lys Gly Pro Asp Gly Thr Asn Val Lys Lys Ser
130                 135                 140                 145

cat aaa gca act gtt cat att cag gtg aac gac gtg aat gag tac gcg   718
His Lys Ala Thr Val His Ile Gln Val Asn Asp Val Asn Glu Tyr Ala
                150                 155                 160

ccc gtg ttc aag gag aag tcc tac aaa gcc acg gtc atc gag ggg aag   766
Pro Val Phe Lys Glu Lys Ser Tyr Lys Ala Thr Val Ile Glu Gly Lys
            165                 170                 175

cag tac gac agc att ttg agg gtg gag gcc gtg gat gcc gac tgc tcc   814
Gln Tyr Asp Ser Ile Leu Arg Val Glu Ala Val Asp Ala Asp Cys Ser
        180                 185                 190

cct cag ttc agc cag att tgc agc tac gaa atc atc act cca gac gtg   862
Pro Gln Phe Ser Gln Ile Cys Ser Tyr Glu Ile Ile Thr Pro Asp Val
    195                 200                 205

ccc ttt act gtt gac aaa gat ggt tat ata aaa aac aca gag aaa tta   910
Pro Phe Thr Val Asp Lys Asp Gly Tyr Ile Lys Asn Thr Glu Lys Leu
210                 215                 220                 225

aac tac ggg aaa gaa cat caa tat aag ctg acc gtc act gcc tat gac   958
Asn Tyr Gly Lys Glu His Gln Tyr Lys Leu Thr Val Thr Ala Tyr Asp
```

```
                230                 235                 240

tgt ggg aag aaa aga gcc aca gaa gat gtt ttg gtg aag atc agc att   1006
Cys Gly Lys Lys Arg Ala Thr Glu Asp Val Leu Val Lys Ile Ser Ile
            245                 250                 255

aag ccc acc tgc acc cct ggg tgg caa gga tgg aac aac agg att gag   1054
Lys Pro Thr Cys Thr Pro Gly Trp Gln Gly Trp Asn Asn Arg Ile Glu
        260                 265                 270

tat gag ccg ggc acc ggc gcg ttg gcc gtc ttt cca aat atc cac ctg   1102
Tyr Glu Pro Gly Thr Gly Ala Leu Ala Val Phe Pro Asn Ile His Leu
    275                 280                 285

gag aca tgt gac gag cca gtc gcc tca gta cag gcc aca gtg gag cta   1150
Glu Thr Cys Asp Glu Pro Val Ala Ser Val Gln Ala Thr Val Glu Leu
290                 295                 300                 305

gaa acc agc cac ata ggg aaa ggc tgc gac cga gac acc tac tca gag   1198
Glu Thr Ser His Ile Gly Lys Gly Cys Asp Arg Asp Thr Tyr Ser Glu
                310                 315                 320

aag tcc ctc cac cgg ctc tgt ggt gcg gcc gcg ggc act gcc gag ctg   1246
Lys Ser Leu His Arg Leu Cys Gly Ala Ala Ala Gly Thr Ala Glu Leu
            325                 330                 335

ctg cca tcc ccg agt gga tcc ctc aac tgg acc atg ggc ctg ccc acc   1294
Leu Pro Ser Pro Ser Gly Ser Leu Asn Trp Thr Met Gly Leu Pro Thr
        340                 345                 350

gac aat ggc cac gac agc gac cag gtg ttt gag ttc aac ggc acc cag   1342
Asp Asn Gly His Asp Ser Asp Gln Val Phe Glu Phe Asn Gly Thr Gln
    355                 360                 365

gca gtg agg atc ccg gat ggc gtc gtg tcg gtc agc ccc aaa gag ccg   1390
Ala Val Arg Ile Pro Asp Gly Val Val Ser Val Ser Pro Lys Glu Pro
370                 375                 380                 385

ttc acc atc tcg gtg tgg atg aga cat ggg cca ttc ggc agg aag aag   1438
Phe Thr Ile Ser Val Trp Met Arg His Gly Pro Phe Gly Arg Lys Lys
                390                 395                 400

gag aca att ctt tgc agt tct gat aaa aca gat atg aat cgg cac cac   1486
Glu Thr Ile Leu Cys Ser Ser Asp Lys Thr Asp Met Asn Arg His His
            405                 410                 415

tac tcc ctc tat gtc cac ggg tgc cgg ctg atc ttc ctc ttc cgt cag   1534
Tyr Ser Leu Tyr Val His Gly Cys Arg Leu Ile Phe Leu Phe Arg Gln
```

```
        420                 425                 430

gat cct tct gag gag aag aaa tac aga cct gca gag ttc cac tgg aag   1582
Asp Pro Ser Glu Glu Lys Lys Tyr Arg Pro Ala Glu Phe His Trp Lys
    435                 440                 445

ttg aat cag gtc tgt gat gag gaa tgg cac cac tac gtc ctc aat gta   1630
Leu Asn Gln Val Cys Asp Glu Glu Trp His His Tyr Val Leu Asn Val
450                 455                 460                 465

gaa ttc ccg agt gtg act ctc tat gtg gat ggc acg tcc cac gag ccc   1678
Glu Phe Pro Ser Val Thr Leu Tyr Val Asp Gly Thr Ser His Glu Pro
                470                 475                 480

ttc tct gtg act gag gat tac ccg ctc cat cca tcc aag ata gaa act   1726
Phe Ser Val Thr Glu Asp Tyr Pro Leu His Pro Ser Lys Ile Glu Thr
            485                 490                 495

cag ctc gtg gtg ggg gct tgc tgg caa gag ttt tca gga gtt gaa aat   1774
Gln Leu Val Val Gly Ala Cys Trp Gln Glu Phe Ser Gly Val Glu Asn
        500                 505                 510

gac aat gaa act gag cct gtg act gtg gcc tct gca ggt ggc gac ctg   1822
Asp Asn Glu Thr Glu Pro Val Thr Val Ala Ser Ala Gly Gly Asp Leu
    515                 520                 525

cac atg acc cag ttt ttc cga ggc aat ctg gct ggc tta act ctc cgt   1870
His Met Thr Gln Phe Phe Arg Gly Asn Leu Ala Gly Leu Thr Leu Arg
530                 535                 540                 545

tcc ggg aaa ctc gcg gat aag aag gtg atc gac tgt ctg tat acc tgc   1918
Ser Gly Lys Leu Ala Asp Lys Lys Val Ile Asp Cys Leu Tyr Thr Cys
                550                 555                 560

aag gag ggg ctg gac ctg cag gtc ctc gaa gac agt ggc aga ggc gtg   1966
Lys Glu Gly Leu Asp Leu Gln Val Leu Glu Asp Ser Gly Arg Gly Val
            565                 570                 575

cag atc caa gca cac ccc agc cag ttg gta ttg acc ttg gag gga gaa   2014
Gln Ile Gln Ala His Pro Ser Gln Leu Val Leu Thr Leu Glu Gly Glu
        580                 585                 590

gac ctc ggg gaa ttg gat aag gcc atg cag cac atc tcg tac ctg aac   2062
Asp Leu Gly Glu Leu Asp Lys Ala Met Gln His Ile Ser Tyr Leu Asn
    595                 600                 605

tcc cgg cag ttc ccc acg ccc gga att cgc aga ctc aaa atc acc agc   2110
Ser Arg Gln Phe Pro Thr Pro Gly Ile Arg Arg Leu Lys Ile Thr Ser
```

```
610                 615                 620                 625

aca atc aag tgt ttt aac gag gcc acc tgc att tcg gtc ccc ccg gta   2158
Thr Ile Lys Cys Phe Asn Glu Ala Thr Cys Ile Ser Val Pro Pro Val
                630                 635                 640

gat ggc tac gtg atg gtt tta cag ccc gag gag ccc aag atc agc ctg   2206
Asp Gly Tyr Val Met Val Leu Gln Pro Glu Glu Pro Lys Ile Ser Leu
            645                 650                 655

agt ggc gtc cac cat ttt gcc cga gca gct tct gaa ttt gaa agc tca   2254
Ser Gly Val His His Phe Ala Arg Ala Ala Ser Glu Phe Glu Ser Ser
        660                 665                 670

gaa ggg gtg ttc ctt ttc cct gag ctt cgc atc atc agc acc atc acg   2302
Glu Gly Val Phe Leu Phe Pro Glu Leu Arg Ile Ile Ser Thr Ile Thr
    675                 680                 685

aga gaa gtg gag cct gaa ggg gac ggg gct gag gac ccc aca gtt caa   2350
Arg Glu Val Glu Pro Glu Gly Asp Gly Ala Glu Asp Pro Thr Val Gln
690                 695                 700                 705

gaa tca ctg gtg tcc gag gag atc gtg cac gac ctg gat acc tgt gag   2398
Glu Ser Leu Val Ser Glu Glu Ile Val His Asp Leu Asp Thr Cys Glu
                710                 715                 720

gtc acg gtg gag gga gag gag ctg aac cac gag cag gag agc ctg gag   2446
Val Thr Val Glu Gly Glu Glu Leu Asn His Glu Gln Glu Ser Leu Glu
            725                 730                 735

gtg gac atg gcc cgc ctg cag cag aag ggc att gaa gtg agc agc tct   2494
Val Asp Met Ala Arg Leu Gln Gln Lys Gly Ile Glu Val Ser Ser Ser
        740                 745                 750

gaa ctg ggc atg acc ttc aca ggc gtg gac acc atg gcc agc tac gag   2542
Glu Leu Gly Met Thr Phe Thr Gly Val Asp Thr Met Ala Ser Tyr Glu
    755                 760                 765

gag gtt ttg cac ctg ctg cgc tat cgg aac tgg cat gcc agg tcc ttg   2590
Glu Val Leu His Leu Leu Arg Tyr Arg Asn Trp His Ala Arg Ser Leu
770                 775                 780                 785

ctt gac cgg aag ttt aag ctc atc tgc tca gag ctg aat ggc cgc tac   2638
Leu Asp Arg Lys Phe Lys Leu Ile Cys Ser Glu Leu Asn Gly Arg Tyr
                790                 795                 800

atc agc aac gaa ttt aag gtg gag gtg aat gta atc cac acg gcc aac   2686
Ile Ser Asn Glu Phe Lys Val Glu Val Asn Val Ile His Thr Ala Asn
```

```
          805                 810                 815

ccc atg gaa cac gcc aac cac atg gct gcc cag cca cag ttc gtg cac   2734
Pro Met Glu His Ala Asn His Met Ala Ala Gln Pro Gln Phe Val His
        820                 825                 830

ccg gaa cac cgc tcc ttt gtt gac ctg tca ggc cac aac ctg gcc aac   2782
Pro Glu His Arg Ser Phe Val Asp Leu Ser Gly His Asn Leu Ala Asn
    835                 840                 845

ccc cac ccg ttc gca gtc gtc ccc agc act gcg aca gtt gtg atc gtg   2830
Pro His Pro Phe Ala Val Val Pro Ser Thr Ala Thr Val Val Ile Val
850                 855                 860                 865

gtg tgc gtc agc ttc ctg gtg ttc atg att atc ctg ggg gta ttt cgg   2878
Val Cys Val Ser Phe Leu Val Phe Met Ile Ile Leu Gly Val Phe Arg
                870                 875                 880

atc cgg gcc gca cat cgg cgg acc atg cgg gat cag gac acc ggg aag   2926
Ile Arg Ala Ala His Arg Arg Thr Met Arg Asp Gln Asp Thr Gly Lys
            885                 890                 895

gag aac gag atg gac tgg gac gac tct gcc ctg acc atc acc gtc aac   2974
Glu Asn Glu Met Asp Trp Asp Asp Ser Ala Leu Thr Ile Thr Val Asn
        900                 905                 910

ccc atg gag acc tat gag gac cag cac agc agt gag gag gag gag gaa   3022
Pro Met Glu Thr Tyr Glu Asp Gln His Ser Ser Glu Glu Glu Glu Glu
    915                 920                 925

gag gaa gag gaa gag gaa agc gag gac ggc gaa gaa gag gat gac atc   3070
Glu Glu Glu Glu Glu Glu Ser Glu Asp Gly Glu Glu Glu Asp Asp Ile
930                 935                 940                 945

acc agc gcc gag tcg gag agc agc gag gag gag gag ggg gag cag ggc   3118
Thr Ser Ala Glu Ser Glu Ser Ser Glu Glu Glu Glu Gly Glu Gln Gly
                950                 955                 960

gac ccc cag aac gca acc cgg cag cag cag ctg gag tgg gat gac tcc   3166
Asp Pro Gln Asn Ala Thr Arg Gln Gln Gln Leu Glu Trp Asp Asp Ser
            965                 970                 975

acc ctc agc tac tgacccgtgc ccccggccac ctcggtttct gctttcgaag       3218
Thr Leu Ser Tyr
        980

actctgctgc catccgttct cccagtccca agggtccacg atgtacaaag tcatttcggc 3278
```

cagtaggtgt gcagacccct ccccgccacg atcgtcgctg tgcttggtgt gtaggaccct 3338

aggctccccg cccaccctct gcctggtcgc gctcttcagt cccacgagga gctgacacgt 3398

cctctctggc cgccatccgg ctcgcacagg ggcctcccag cgcctcaggc cccgcgtttg 3458

tgtctggagt ctccccccgg ggagaggaca ctggcccctc gcactccaga aaagccatgc 3518

cagctgggct cgttgacaaa gggtaaaaca tgctcatccc acccggtaat tcatttttt 3578

ctttttttta aaaaaagttt ttatttttt ccaaactagt gcatgtatta aataatgggc 3638

agggggggg gtatgtttta ggatgattaa ctggactttt taatattttt tgtaaaataa 3698

aa 3700

<210> 2
<211> 981
<212> PRT
<213> Homo sapiens

<400> 2

Met Leu Arg Arg Pro Ala Pro Glu Leu Ala Pro Ala Ala Arg Leu Leu
1 5 10 15

Leu Ala Gly Leu Leu Cys Gly Gly Gly Val Trp Ala Ala Arg Val Asn
20 25 30

Lys His Lys Pro Trp Leu Glu Pro Thr Tyr His Gly Ile Val Thr Glu
35 40 45

Asn Asp Asn Thr Val Leu Leu Asp Pro Pro Leu Ile Ala Leu Asp Lys
50 55 60

Asp Ala Pro Leu Arg Phe Ala Glu Ser Phe Glu Val Thr Val Thr Lys
65 70 75 80

Glu Gly Glu Ile Cys Gly Phe Lys Ile His Gly Gln Asn Val Pro Phe
85 90 95

Asp Ala Val Val Val Asp Lys Ser Thr Gly Glu Gly Val Ile Arg Ser
100 105 110

Lys Glu Lys Leu Asp Cys Glu Leu Gln Lys Asp Tyr Ser Phe Thr Ile
115 120 125

Gln Ala Tyr Asp Cys Gly Lys Gly Pro Asp Gly Thr Asn Val Lys Lys

130 135 140

Ser His Lys Ala Thr Val His Ile Gln Val Asn Asp Val Asn Glu Tyr
145 150 155 160

Ala Pro Val Phe Lys Glu Lys Ser Tyr Lys Ala Thr Val Ile Glu Gly
165 170 175

Lys Gln Tyr Asp Ser Ile Leu Arg Val Glu Ala Val Asp Ala Asp Cys
180 185 190

Ser Pro Gln Phe Ser Gln Ile Cys Ser Tyr Glu Ile Ile Thr Pro Asp
195 200 205

Val Pro Phe Thr Val Asp Lys Asp Gly Tyr Ile Lys Asn Thr Glu Lys
210 215 220

Leu Asn Tyr Gly Lys Glu His Gln Tyr Lys Leu Thr Val Thr Ala Tyr
225 230 235 240

Asp Cys Gly Lys Lys Arg Ala Thr Glu Asp Val Leu Val Lys Ile Ser
245 250 255

Ile Lys Pro Thr Cys Thr Pro Gly Trp Gln Gly Trp Asn Asn Arg Ile
260 265 270

Glu Tyr Glu Pro Gly Thr Gly Ala Leu Ala Val Phe Pro Asn Ile His
275 280 285

Leu Glu Thr Cys Asp Glu Pro Val Ala Ser Val Gln Ala Thr Val Glu
290 295 300

Leu Glu Thr Ser His Ile Gly Lys Gly Cys Asp Arg Asp Thr Tyr Ser
305 310 315 320

Glu Lys Ser Leu His Arg Leu Cys Gly Ala Ala Ala Gly Thr Ala Glu
325 330 335

Leu Leu Pro Ser Pro Ser Gly Ser Leu Asn Trp Thr Met Gly Leu Pro
340 345 350

Thr Asp Asn Gly His Asp Ser Asp Gln Val Phe Glu Phe Asn Gly Thr
355 360 365

Gln Ala Val Arg Ile Pro Asp Gly Val Val Ser Val Ser Pro Lys Glu
370 375 380

Pro Phe Thr Ile Ser Val Trp Met Arg His Gly Pro Phe Gly Arg Lys

385 390 395 400

Lys Glu Thr Ile Leu Cys Ser Ser Asp Lys Thr Asp Met Asn Arg His
405 410 415

His Tyr Ser Leu Tyr Val His Gly Cys Arg Leu Ile Phe Leu Phe Arg
420 425 430

Gln Asp Pro Ser Glu Glu Lys Lys Tyr Arg Pro Ala Glu Phe His Trp
435 440 445

Lys Leu Asn Gln Val Cys Asp Glu Glu Trp His His Tyr Val Leu Asn
450 455 460

Val Glu Phe Pro Ser Val Thr Leu Tyr Val Asp Gly Thr Ser His Glu
465 470 475 480

Pro Phe Ser Val Thr Glu Asp Tyr Pro Leu His Pro Ser Lys Ile Glu
485 490 495

Thr Gln Leu Val Val Gly Ala Cys Trp Gln Glu Phe Ser Gly Val Glu
500 505 510

Asn Asp Asn Glu Thr Glu Pro Val Thr Val Ala Ser Ala Gly Gly Asp
515 520 525

Leu His Met Thr Gln Phe Phe Arg Gly Asn Leu Ala Gly Leu Thr Leu
530 535 540

Arg Ser Gly Lys Leu Ala Asp Lys Lys Val Ile Asp Cys Leu Tyr Thr
545 550 555 560

Cys Lys Glu Gly Leu Asp Leu Gln Val Leu Glu Asp Ser Gly Arg Gly
565 570 575

Val Gln Ile Gln Ala His Pro Ser Gln Leu Val Leu Thr Leu Glu Gly
580 585 590

Glu Asp Leu Gly Glu Leu Asp Lys Ala Met Gln His Ile Ser Tyr Leu
595 600 605

Asn Ser Arg Gln Phe Pro Thr Pro Gly Ile Arg Arg Leu Lys Ile Thr
610 615 620

Ser Thr Ile Lys Cys Phe Asn Glu Ala Thr Cys Ile Ser Val Pro Pro
625 630 635 640

Val Asp Gly Tyr Val Met Val Leu Gln Pro Glu Glu Pro Lys Ile Ser

645 650 655

Leu Ser Gly Val His His Phe Ala Arg Ala Ala Ser Glu Phe Glu Ser
660 665 670

Ser Glu Gly Val Phe Leu Phe Pro Glu Leu Arg Ile Ile Ser Thr Ile
675 680 685

Thr Arg Glu Val Glu Pro Glu Gly Asp Gly Ala Glu Asp Pro Thr Val
690 695 700

Gln Glu Ser Leu Val Ser Glu Glu Ile Val His Asp Leu Asp Thr Cys
705 710 715 720

Glu Val Thr Val Glu Gly Glu Glu Leu Asn His Glu Gln Glu Ser Leu
725 730 735

Glu Val Asp Met Ala Arg Leu Gln Gln Lys Gly Ile Glu Val Ser Ser
740 745 750

Ser Glu Leu Gly Met Thr Phe Thr Gly Val Asp Thr Met Ala Ser Tyr
755 760 765

Glu Glu Val Leu His Leu Leu Arg Tyr Arg Asn Trp His Ala Arg Ser
770 775 780

Leu Leu Asp Arg Lys Phe Lys Leu Ile Cys Ser Glu Leu Asn Gly Arg
785 790 795 800

Tyr Ile Ser Asn Glu Phe Lys Val Glu Val Asn Val Ile His Thr Ala
805 810 815

Asn Pro Met Glu His Ala Asn His Met Ala Ala Gln Pro Gln Phe Val
820 825 830

His Pro Glu His Arg Ser Phe Val Asp Leu Ser Gly His Asn Leu Ala
835 840 845

Asn Pro His Pro Phe Ala Val Val Pro Ser Thr Ala Thr Val Val Ile
850 855 860

Val Val Cys Val Ser Phe Leu Val Phe Met Ile Ile Leu Gly Val Phe
865 870 875 880

Arg Ile Arg Ala Ala His Arg Arg Thr Met Arg Asp Gln Asp Thr Gly
885 890 895

Lys Glu Asn Glu Met Asp Trp Asp Asp Ser Ala Leu Thr Ile Thr Val

900 905 910

Asn Pro Met Glu Thr Tyr Glu Asp Gln His Ser Ser Glu Glu Glu Glu
915 920 925

Glu Glu Glu Glu Glu Glu Glu Ser Glu Asp Gly Glu Glu Glu Asp Asp
930 935 940

Ile Thr Ser Ala Glu Ser Glu Ser Ser Glu Glu Glu Glu Gly Glu Gln
945 950 955 960

Gly Asp Pro Gln Asn Ala Thr Arg Gln Gln Gln Leu Glu Trp Asp Asp
965 970 975

Ser Thr Leu Ser Tyr
980

<210> 3
<211> 4421
<212> DNA
<213> Homo sapiens

<220>
<221> 5'UTR
<222> (1)..(10)

<220>
<221> CDS
<222> (11)..(2875)

<220>
<221> 3'UTR
<222> (2876)..(4421)

<400> 3
tgctgcgagg atg ctg cct ggg cgg ctg tgc tgg gtg ccg ctc ctg ctg 49
Met Leu Pro Gly Arg Leu Cys Trp Val Pro Leu Leu Leu
1 5 10

gcg ctg ggc gtg ggg agc ggc agc ggc ggt ggc ggg gac agc cgg cag 97
Ala Leu Gly Val Gly Ser Gly Ser Gly Gly Gly Gly Asp Ser Arg Gln
15 20 25

cgc cgc ctc ctc gcg gct aaa gtc aat aag cac aag cca tgg atc gag 145
Arg Arg Leu Leu Ala Ala Lys Val Asn Lys His Lys Pro Trp Ile Glu
30 35 40 45

```
act tca tat cat gga gtc ata act gag aac aat gac aca gtc att ttg   193
Thr Ser Tyr His Gly Val Ile Thr Glu Asn Asn Asp Thr Val Ile Leu
                 50                  55                  60

gac cca cca ctg gta gcc ctg gat aaa gat gca ccg gtt cct ttt gca   241
Asp Pro Pro Leu Val Ala Leu Asp Lys Asp Ala Pro Val Pro Phe Ala
            65                  70                  75

ggg gaa atc tgt gcg ttc aag atc cat ggc cag gag ctg ccc ttt gag   289
Gly Glu Ile Cys Ala Phe Lys Ile His Gly Gln Glu Leu Pro Phe Glu
         80                  85                  90

gct gtg gtg ctc aac aag aca tca gga gag ggc cgg ctc cgt gcc aag   337
Ala Val Val Leu Asn Lys Thr Ser Gly Glu Gly Arg Leu Arg Ala Lys
     95                 100                 105

agc ccc att gac tgt gag ttg cag aag gag tac aca ttc atc atc cag   385
Ser Pro Ile Asp Cys Glu Leu Gln Lys Glu Tyr Thr Phe Ile Ile Gln
110                 115                 120                 125

gcc tat gac tgt ggt gct ggg ccc cac gag aca gcc tgg aaa aag tca   433
Ala Tyr Asp Cys Gly Ala Gly Pro His Glu Thr Ala Trp Lys Lys Ser
                130                 135                 140

cac aag gcc gtg gtc cat ata cag gtg aag gat gtc aac gag ttt gct   481
His Lys Ala Val Val His Ile Gln Val Lys Asp Val Asn Glu Phe Ala
            145                 150                 155

ccc acc ttc aaa gag cca gcc tac aag gct gtt gtg acg gag ggc aag   529
Pro Thr Phe Lys Glu Pro Ala Tyr Lys Ala Val Val Thr Glu Gly Lys
        160                 165                 170

atc tat gac agc att ctg cag gtg gag gcc att gac gag gac tgc tcc   577
Ile Tyr Asp Ser Ile Leu Gln Val Glu Ala Ile Asp Glu Asp Cys Ser
    175                 180                 185

cca cag tac agc cag atc tgc aac tat gaa atc gtc acc aca gat gtg   625
Pro Gln Tyr Ser Gln Ile Cys Asn Tyr Glu Ile Val Thr Thr Asp Val
190                 195                 200                 205

cct ttt gcc atc gac aga aat ggc aac atc agg aac act gag aag ctg   673
Pro Phe Ala Ile Asp Arg Asn Gly Asn Ile Arg Asn Thr Glu Lys Leu
                210                 215                 220

agc tat gac aaa caa cac cag tat gag atc ctg gtg acc gcc tac gac   721
Ser Tyr Asp Lys Gln His Gln Tyr Glu Ile Leu Val Thr Ala Tyr Asp
            225                 230                 235
```

```
tgt gga cag aag ccc gct gct cag gac acc ctg gtg cag gtg gat gtg   769
Cys Gly Gln Lys Pro Ala Ala Gln Asp Thr Leu Val Gln Val Asp Val
        240                 245                 250

aag cca gtt tgc aag cct ggc tgg caa gac tgg acc aag agg att gag   817
Lys Pro Val Cys Lys Pro Gly Trp Gln Asp Trp Thr Lys Arg Ile Glu
    255                 260                 265

tac cag cct ggc tcc ggg agc atg ccc ctg ttc ccc agc atc cac ctg   865
Tyr Gln Pro Gly Ser Gly Ser Met Pro Leu Phe Pro Ser Ile His Leu
270                 275                 280                 285

gag acg tgc gat gga gcc gtg tct tcc ctc cag atc gtc aca gag ctg   913
Glu Thr Cys Asp Gly Ala Val Ser Ser Leu Gln Ile Val Thr Glu Leu
                290                 295                 300

cag act aat tac att ggg aag ggt tgt gac cgg gag acc tac tct gag   961
Gln Thr Asn Tyr Ile Gly Lys Gly Cys Asp Arg Glu Thr Tyr Ser Glu
            305                 310                 315

aaa tcc ctt cag aag tta tgt gga gcc tcc tct ggc atc att gac ctc   1009
Lys Ser Leu Gln Lys Leu Cys Gly Ala Ser Ser Gly Ile Ile Asp Leu
        320                 325                 330

ttg cca tcc cct agc gct gcc acc aac tgg act gca gga ctg ctg gtg   1057
Leu Pro Ser Pro Ser Ala Ala Thr Asn Trp Thr Ala Gly Leu Leu Val
    335                 340                 345

gac agc agt gag atg atc ttc aag ttt gac ggc agg cag ggt gcc aaa   1105
Asp Ser Ser Glu Met Ile Phe Lys Phe Asp Gly Arg Gln Gly Ala Lys
350                 355                 360                 365

atc ccc gat ggg att gtg ccc aag aac ctg acc gat cag ttc acc atc   1153
Ile Pro Asp Gly Ile Val Pro Lys Asn Leu Thr Asp Gln Phe Thr Ile
                370                 375                 380

acc atg tgg atg aaa cac ggc ccc agc cct ggt gtg aga gcc gag aag   1201
Thr Met Trp Met Lys His Gly Pro Ser Pro Gly Val Arg Ala Glu Lys
            385                 390                 395

gaa acc atc ctc tgc aac tca gac aaa acc gaa atg aac cgg cat cac   1249
Glu Thr Ile Leu Cys Asn Ser Asp Lys Thr Glu Met Asn Arg His His
        400                 405                 410

tat gcc ctg tat gtg cac aac tgc cgc ctc gtc ttt ctc ttg cgg aag   1297
Tyr Ala Leu Tyr Val His Asn Cys Arg Leu Val Phe Leu Leu Arg Lys
    415                 420                 425
```

```
gac ttc gac cag gct gac acc ttt cgc ccc gcg gag ttc cac tgg aag   1345
Asp Phe Asp Gln Ala Asp Thr Phe Arg Pro Ala Glu Phe His Trp Lys
430                 435                 440                 445

ctg gat cag att tgt gac aaa gag tgg cac tac tat gtc atc aat gtg   1393
Leu Asp Gln Ile Cys Asp Lys Glu Trp His Tyr Tyr Val Ile Asn Val
                450                 455                 460

gag ttt cct gtg gta acc tta tac atg gat gga gca aca tat gaa cca   1441
Glu Phe Pro Val Val Thr Leu Tyr Met Asp Gly Ala Thr Tyr Glu Pro
            465                 470                 475

tac ctg gtg acc aac gac tgg ccc att cat cca tct cac ata gcc atg   1489
Tyr Leu Val Thr Asn Asp Trp Pro Ile His Pro Ser His Ile Ala Met
        480                 485                 490

caa ctc aca gtc ggc gct tgt tgg caa gga gga gaa gtc acc aaa cca   1537
Gln Leu Thr Val Gly Ala Cys Trp Gln Gly Gly Glu Val Thr Lys Pro
    495                 500                 505

cag ttt gct cag ttc ttt cat gga agc ctg gcc agt ctc acc atc cgc   1585
Gln Phe Ala Gln Phe Phe His Gly Ser Leu Ala Ser Leu Thr Ile Arg
510                 515                 520                 525

cct ggc aaa atg gaa agc cag aag gtg atc tcc tgc ctg cag gcc tgc   1633
Pro Gly Lys Met Glu Ser Gln Lys Val Ile Ser Cys Leu Gln Ala Cys
                530                 535                 540

aag gaa ggg ctg gac att aat tcc ttg gaa agc ctt ggc caa gga ata   1681
Lys Glu Gly Leu Asp Ile Asn Ser Leu Glu Ser Leu Gly Gln Gly Ile
            545                 550                 555

aag tat cac ttc aac ccc tcg cag tcc atc ctg gtg atg gaa ggt gac   1729
Lys Tyr His Phe Asn Pro Ser Gln Ser Ile Leu Val Met Glu Gly Asp
        560                 565                 570

gac att ggg aac att aac cgt gct ctc cag aaa gtc tcc tac atc aac   1777
Asp Ile Gly Asn Ile Asn Arg Ala Leu Gln Lys Val Ser Tyr Ile Asn
    575                 580                 585

tcc agg cag ttc cca acg gcg ggt gtg cgg cgc ctc aaa gta tcc tcc   1825
Ser Arg Gln Phe Pro Thr Ala Gly Val Arg Arg Leu Lys Val Ser Ser
590                 595                 600                 605

aaa gtc cag tgc ttt ggg gaa gac gta tgc atc agt atc cct gag gta   1873
Lys Val Gln Cys Phe Gly Glu Asp Val Cys Ile Ser Ile Pro Glu Val
                610                 615                 620
```

```
gat gcc tat gtg atg gtc ctc cag gcc atc gag ccc cgg atc acc ctc   1921
Asp Ala Tyr Val Met Val Leu Gln Ala Ile Glu Pro Arg Ile Thr Leu
            625                 630                 635

cgg ggc aca gac cac ttc tgg aga cct gct gcc cag ttt gaa agt gcc   1969
Arg Gly Thr Asp His Phe Trp Arg Pro Ala Ala Gln Phe Glu Ser Ala
        640                 645                 650

agg gga gtg acc ctc ttc cct gat atc aag att gtg agc acc ttc gcc   2017
Arg Gly Val Thr Leu Phe Pro Asp Ile Lys Ile Val Ser Thr Phe Ala
    655                 660                 665

aaa acc gaa gcc ccc ggg gac gtg aaa acc aca gac ccc aaa tca gaa   2065
Lys Thr Glu Ala Pro Gly Asp Val Lys Thr Thr Asp Pro Lys Ser Glu
670                 675                 680                 685

gtc tta gag gaa atg ctt cat aac tta gat ttc tgt gac att ttg gtg   2113
Val Leu Glu Glu Met Leu His Asn Leu Asp Phe Cys Asp Ile Leu Val
                690                 695                 700

atc gga ggg gac ttg gac cca agg cag gag tgc ttg gag ctc aac cac   2161
Ile Gly Gly Asp Leu Asp Pro Arg Gln Glu Cys Leu Glu Leu Asn His
            705                 710                 715

agt gag ctc cac caa cga cac ctg gat gcc act aat tct act gca ggc   2209
Ser Glu Leu His Gln Arg His Leu Asp Ala Thr Asn Ser Thr Ala Gly
        720                 725                 730

tac tcc atc tac ggt gtg ggc tcc atg agc cgc tat gag cag gtg cta   2257
Tyr Ser Ile Tyr Gly Val Gly Ser Met Ser Arg Tyr Glu Gln Val Leu
    735                 740                 745

cat cac atc cgc tac cgc aac tgg cgt ccg gct tcc ctt gag gcc cgg   2305
His His Ile Arg Tyr Arg Asn Trp Arg Pro Ala Ser Leu Glu Ala Arg
750                 755                 760                 765

cgt ttc cgg att aag tgc tca gaa ctc aat ggg cgc tac act agc aat   2353
Arg Phe Arg Ile Lys Cys Ser Glu Leu Asn Gly Arg Tyr Thr Ser Asn
                770                 775                 780

gag ttc aac ttg gag gtc agc atc ctt cat gaa gac caa gtc tca gat   2401
Glu Phe Asn Leu Glu Val Ser Ile Leu His Glu Asp Gln Val Ser Asp
            785                 790                 795

aag gag cat gtc aat cat ctg att gtg cag cct ccc ttc ctc cag tct   2449
Lys Glu His Val Asn His Leu Ile Val Gln Pro Pro Phe Leu Gln Ser
        800                 805                 810
```

```
gtc cat cat cct gag tcc cgg agt agc atc cag cac agt tca gtg gtc   2497
Val His His Pro Glu Ser Arg Ser Ser Ile Gln His Ser Ser Val Val
    815                 820                 825

cca agc att gcc aca gtg gtc atc atc atc tcc gtg tgc atg ctt gtg   2545
Pro Ser Ile Ala Thr Val Val Ile Ile Ile Ser Val Cys Met Leu Val
830                 835                 840                 845

ttt gtc gtg gcc atg ggt gtg tac cgg gtc cgg atc gcc cac cag cac   2593
Phe Val Val Ala Met Gly Val Tyr Arg Val Arg Ile Ala His Gln His
                850                 855                 860

ttc atc cag gag act gag gct gcc aag gaa tct gag atg gac tgg gac   2641
Phe Ile Gln Glu Thr Glu Ala Ala Lys Glu Ser Glu Met Asp Trp Asp
            865                 870                 875

gat tct gcg ctg act atc aca gtc aac ccc atg gag aaa cat gaa gga   2689
Asp Ser Ala Leu Thr Ile Thr Val Asn Pro Met Glu Lys His Glu Gly
        880                 885                 890

cca ggg cat ggg gaa gat gag act gag gga gaa gag gag gaa gaa gcc   2737
Pro Gly His Gly Glu Asp Glu Thr Glu Gly Glu Glu Glu Glu Glu Ala
    895                 900                 905

gag gaa gaa atg agc tcc agc agt ggc tct gac gac agc gaa gag gag   2785
Glu Glu Glu Met Ser Ser Ser Ser Gly Ser Asp Asp Ser Glu Glu Glu
910                 915                 920                 925

gag gag gag gaa ggg atg ggc aga ggc aga cat ggg cag aat gga gcc   2833
Glu Glu Glu Glu Gly Met Gly Arg Gly Arg His Gly Gln Asn Gly Ala
                930                 935                 940

agg caa gcc cag ctg gag tgg gat gac tcc acc ctc ccc tac           2875
Arg Gln Ala Gln Leu Glu Trp Asp Asp Ser Thr Leu Pro Tyr
            945                 950                 955

tagtgcccag gggtctgctg cctggcccac atgtcccttt tgtaaaccct gacccagtgt 2935

atgcccatgt ctatcatacc tcacctctga tgtctgtgac atgtctggga aggccttctc 2995

cagcttcctg gagcccaccc tttaagcctt gggcactccc tgtgtttcat ccatggggaa 3055

gttccaagaa gcccagcatg gccatcagtg aggacttcag ggtagacttt gtcctgtagc 3115

ctccacttct gccctaagtt ccccagcatc ctgactacct gtctgcagag tttgcctttg 3175

ttttttcctg cagggaagaa ggcccacctt tgtgtcactc acctccccag gctcagagtc 3235
```

cccaaggccc tggggttcca actcactgtg cgtctcctcc acacagacca gtaggttctc 3295

ctatgctgac tccaggttgc ttcatacaag gagggtggtt gaacttcaca cacgtaaggt 3355

cttagtgctt aacagtttaa aggaaagtcc ttgttgaggc agaactaagt ttacagggaa 3415

aggtacacac attctctctc tctctctctc tctgtctatc tagttcccca gcttggagag 3475

cctttcccct tgcttctttc tgaggccata taagcttata agaaaagtcc caaaccaaga 3535

ataggtcctt ggccacaagc agggtctgat cccccatcag agctatctga gcctgcctgt 3595

ctgggcacct gctgcaacca tgcagctacc ctgccagggg cactcagcaa acagaaccac 3655

agggcccagg aggcattcca cacaggcact gccccaggac aacacaacaa ggacagtcac 3715

aacaaggaca acaaggacac aacacaacac acaacaagga cagtcacaac aagcctagag 3775

ccagaaagca gatggaaatg ctaatgaggt caaacgtagg cttcatggtg ggtggagtgg 3835

gggtggctgg gctcccccag gacagagggg accctgaggt tggcaaggct ctcaccactc 3895

agccttatgg tcccttatct cctatcttcc ctcttgagaa aatacacgct ttctgcatgt 3955

attagaaacg cacgagctcc accaagtcta caatgaaagt ttgaaattta actgcaagga 4015

attagaagca tatttgcaat cattgcagct tcttctttct tctgctcata aaaggaggaa 4075

cactttagat agagggcaaa tatatctgaa aacctaattt ctttcttttt ttgataagga 4135

aatcttttcc atctccatcc taacatgcac aacctgtgaa gagaattgtt tctatagtaa 4195

ctggtctgtg atcttttgtg gccaagagaa tagcaggcaa gaattagggc cttgacagaa 4255

tttccacgaa gctctgagaa catgtttgtt tcgaatgtct gattcctctt tgtcatcaat 4315

gtgtatgctc tgtccccatc cttcactcct cctcaagctc acaccaattg gtttggcaca 4375

ggcacagagc tggtccctag ttaagtggca tttatgttaa aaaaaa 4421

<210> 4
<211> 955
<212> PRT
<213> Homo sapiens

<400> 4

```
Met Leu Pro Gly Arg Leu Cys Trp Val Pro Leu Leu Leu Ala Leu Gly
 1               5                  10                  15

Val Gly Ser Gly Ser Gly Gly Gly Gly Asp Ser Arg Gln Arg Arg Leu
            20                  25                  30

Leu Ala Ala Lys Val Asn Lys His Lys Pro Trp Ile Glu Thr Ser Tyr
        35                  40                  45

His Gly Val Ile Thr Glu Asn Asn Asp Thr Val Ile Leu Asp Pro Pro
    50                  55                  60

Leu Val Ala Leu Asp Lys Asp Ala Pro Val Pro Phe Ala Gly Glu Ile
65                  70                  75                  80

Cys Ala Phe Lys Ile His Gly Gln Glu Leu Pro Phe Glu Ala Val Val
                85                  90                  95

Leu Asn Lys Thr Ser Gly Glu Gly Arg Leu Arg Ala Lys Ser Pro Ile
            100                 105                 110

Asp Cys Glu Leu Gln Lys Glu Tyr Thr Phe Ile Ile Gln Ala Tyr Asp
        115                 120                 125

Cys Gly Ala Gly Pro His Glu Thr Ala Trp Lys Lys Ser His Lys Ala
    130                 135                 140

Val Val His Ile Gln Val Lys Asp Val Asn Glu Phe Ala Pro Thr Phe
145                 150                 155                 160

Lys Glu Pro Ala Tyr Lys Ala Val Val Thr Glu Gly Lys Ile Tyr Asp
                165                 170                 175

Ser Ile Leu Gln Val Glu Ala Ile Asp Glu Asp Cys Ser Pro Gln Tyr
            180                 185                 190

Ser Gln Ile Cys Asn Tyr Glu Ile Val Thr Thr Asp Val Pro Phe Ala
        195                 200                 205

Ile Asp Arg Asn Gly Asn Ile Arg Asn Thr Glu Lys Leu Ser Tyr Asp
    210                 215                 220

Lys Gln His Gln Tyr Glu Ile Leu Val Thr Ala Tyr Asp Cys Gly Gln
225                 230                 235                 240

Lys Pro Ala Ala Gln Asp Thr Leu Val Gln Val Asp Val Lys Pro Val
                245                 250                 255
```

```
Cys Lys Pro Gly Trp Gln Asp Trp Thr Lys Arg Ile Glu Tyr Gln Pro
            260                 265                 270

Gly Ser Gly Ser Met Pro Leu Phe Pro Ser Ile His Leu Glu Thr Cys
        275                 280                 285

Asp Gly Ala Val Ser Ser Leu Gln Ile Val Thr Glu Leu Gln Thr Asn
    290                 295                 300

Tyr Ile Gly Lys Gly Cys Asp Arg Glu Thr Tyr Ser Glu Lys Ser Leu
305                 310                 315                 320

Gln Lys Leu Cys Gly Ala Ser Ser Gly Ile Ile Asp Leu Leu Pro Ser
                325                 330                 335

Pro Ser Ala Ala Thr Asn Trp Thr Ala Gly Leu Leu Val Asp Ser Ser
            340                 345                 350

Glu Met Ile Phe Lys Phe Asp Gly Arg Gln Gly Ala Lys Ile Pro Asp
        355                 360                 365

Gly Ile Val Pro Lys Asn Leu Thr Asp Gln Phe Thr Ile Thr Met Trp
    370                 375                 380

Met Lys His Gly Pro Ser Pro Gly Val Arg Ala Glu Lys Glu Thr Ile
385                 390                 395                 400

Leu Cys Asn Ser Asp Lys Thr Glu Met Asn Arg His His Tyr Ala Leu
                405                 410                 415

Tyr Val His Asn Cys Arg Leu Val Phe Leu Leu Arg Lys Asp Phe Asp
            420                 425                 430

Gln Ala Asp Thr Phe Arg Pro Ala Glu Phe His Trp Lys Leu Asp Gln
        435                 440                 445

Ile Cys Asp Lys Glu Trp His Tyr Tyr Val Ile Asn Val Glu Phe Pro
    450                 455                 460

Val Val Thr Leu Tyr Met Asp Gly Ala Thr Tyr Glu Pro Tyr Leu Val
465                 470                 475                 480

Thr Asn Asp Trp Pro Ile His Pro Ser His Ile Ala Met Gln Leu Thr
                485                 490                 495

Val Gly Ala Cys Trp Gln Gly Gly Glu Val Thr Lys Pro Gln Phe Ala
            500                 505                 510
```

```
Gln Phe Phe His Gly Ser Leu Ala Ser Leu Thr Ile Arg Pro Gly Lys
        515                 520                 525

Met Glu Ser Gln Lys Val Ile Ser Cys Leu Gln Ala Cys Lys Glu Gly
    530                 535                 540

Leu Asp Ile Asn Ser Leu Glu Ser Leu Gly Gln Gly Ile Lys Tyr His
545                 550                 555                 560

Phe Asn Pro Ser Gln Ser Ile Leu Val Met Glu Gly Asp Asp Ile Gly
                565                 570                 575

Asn Ile Asn Arg Ala Leu Gln Lys Val Ser Tyr Ile Asn Ser Arg Gln
            580                 585                 590

Phe Pro Thr Ala Gly Val Arg Arg Leu Lys Val Ser Ser Lys Val Gln
        595                 600                 605

Cys Phe Gly Glu Asp Val Cys Ile Ser Ile Pro Glu Val Asp Ala Tyr
    610                 615                 620

Val Met Val Leu Gln Ala Ile Glu Pro Arg Ile Thr Leu Arg Gly Thr
625                 630                 635                 640

Asp His Phe Trp Arg Pro Ala Ala Gln Phe Glu Ser Ala Arg Gly Val
                645                 650                 655

Thr Leu Phe Pro Asp Ile Lys Ile Val Ser Thr Phe Ala Lys Thr Glu
            660                 665                 670

Ala Pro Gly Asp Val Lys Thr Thr Asp Pro Lys Ser Glu Val Leu Glu
        675                 680                 685

Glu Met Leu His Asn Leu Asp Phe Cys Asp Ile Leu Val Ile Gly Gly
    690                 695                 700

Asp Leu Asp Pro Arg Gln Glu Cys Leu Glu Leu Asn His Ser Glu Leu
705                 710                 715                 720

His Gln Arg His Leu Asp Ala Thr Asn Ser Thr Ala Gly Tyr Ser Ile
                725                 730                 735

Tyr Gly Val Gly Ser Met Ser Arg Tyr Glu Gln Val Leu His His Ile
            740                 745                 750

Arg Tyr Arg Asn Trp Arg Pro Ala Ser Leu Glu Ala Arg Arg Phe Arg
        755                 760                 765
```

```
Ile Lys Cys Ser Glu Leu Asn Gly Arg Tyr Thr Ser Asn Glu Phe Asn
    770                 775                 780

Leu Glu Val Ser Ile Leu His Glu Asp Gln Val Ser Asp Lys Glu His
785                 790                 795                 800

Val Asn His Leu Ile Val Gln Pro Pro Phe Leu Gln Ser Val His His
                805                 810                 815

Pro Glu Ser Arg Ser Ser Ile Gln His Ser Ser Val Val Pro Ser Ile
            820                 825                 830

Ala Thr Val Val Ile Ile Ile Ser Val Cys Met Leu Val Phe Val Val
        835                 840                 845

Ala Met Gly Val Tyr Arg Val Arg Ile Ala His Gln His Phe Ile Gln
    850                 855                 860

Glu Thr Glu Ala Ala Lys Glu Ser Glu Met Asp Trp Asp Asp Ser Ala
865                 870                 875                 880

Leu Thr Ile Thr Val Asn Pro Met Glu Lys His Glu Gly Pro Gly His
                885                 890                 895

Gly Glu Asp Glu Thr Glu Gly Glu Glu Glu Glu Glu Ala Glu Glu Glu
            900                 905                 910

Met Ser Ser Ser Ser Gly Ser Asp Asp Ser Glu Glu Glu Glu Glu Glu
        915                 920                 925

Glu Gly Met Gly Arg Gly Arg His Gly Gln Asn Gly Ala Arg Gln Ala
    930                 935                 940

Gln Leu Glu Trp Asp Asp Ser Thr Leu Pro Tyr
945                 950                 955
```

<210> 5
<211> 3187
<212> DNA
<213> Homo sapiens

<220>
<221> 5'UTR
<222> (1)..(275)

<220>

<221> CDS
<222> (276)..(3143)

<400> 5
ctgcagtagc ggggttgggg tgggagtgag agagtgagga cgctgggctg ggggaaacgg 60

gaagccgctg caagtccacc gcctcagcta cccagattgg gatctgccca ggcccgcttt 120

atggactagt gtgggcggca ggctcctttc cgtccctgcc ctgctgtacc ccgctccttg 180

gagacccccc gtatccctcc cgcaaggtgg aatccgcagg ctggaggctc ccaggggagg 240

caaacgcctg gccctgccct gccccacgcc gcacc atg acc ctc ctg ctg ctg 293
Met Thr Leu Leu Leu Leu
1 5

ccc ctt ctg ctg gcc tct ctg ctc gcg tcc tgc tcc tgt aac aaa gcc 341
Pro Leu Leu Leu Ala Ser Leu Leu Ala Ser Cys Ser Cys Asn Lys Ala
10 15 20

aac aag cac aag cca tgg att gag gca gag tac cag ggc atc gtc atg 389
Asn Lys His Lys Pro Trp Ile Glu Ala Glu Tyr Gln Gly Ile Val Met
25 30 35

gag aat gac aac acg gtc cta ctg aat cca cca ctc ttt gcc ttg gac 437
Glu Asn Asp Asn Thr Val Leu Leu Asn Pro Pro Leu Phe Ala Leu Asp
40 45 50

aag gat gcc ccg ctg cgc tat gca ggt gag atc tgc ggc ttc cgg ctc 485
Lys Asp Ala Pro Leu Arg Tyr Ala Gly Glu Ile Cys Gly Phe Arg Leu
55 60 65 70

cat ggg tct ggg gtg ccc ttt gag gct gtg atc ctt gac aag gcg aca 533
His Gly Ser Gly Val Pro Phe Glu Ala Val Ile Leu Asp Lys Ala Thr
75 80 85

gga gag ggg ctg atc cgg gcc aag gag cct gtg gac tgc gag gcc cag 581
Gly Glu Gly Leu Ile Arg Ala Lys Glu Pro Val Asp Cys Glu Ala Gln
90 95 100

aag gaa cac acc ttc acc atc cag gcc tat gac tgt ggc gag ggc ccc 629
Lys Glu His Thr Phe Thr Ile Gln Ala Tyr Asp Cys Gly Glu Gly Pro
105 110 115

gac ggg gcc aac acc aag aag tcc cac aag gcc act gtg cat gtg cgg 677
Asp Gly Ala Asn Thr Lys Lys Ser His Lys Ala Thr Val His Val Arg
120 125 130

```
gtc aac gat gtg aac gag ttt gcc cca gtg ttt gtg gaa cgg ctg tat   725
Val Asn Asp Val Asn Glu Phe Ala Pro Val Phe Val Glu Arg Leu Tyr
135                 140                 145                 150

cgt gcg gct gtg aca gag ggg aag ctg tac gat cgc atc ctg cgg gtg   773
Arg Ala Ala Val Thr Glu Gly Lys Leu Tyr Asp Arg Ile Leu Arg Val
                155                 160                 165

gaa gcc att gac ggt gac tgc tcc ccc cag tac agc cag atc tgc tac   821
Glu Ala Ile Asp Gly Asp Cys Ser Pro Gln Tyr Ser Gln Ile Cys Tyr
            170                 175                 180

tat gag att ctc aca ccc aac acc cct ttc ctc att gac aat gac ggg   869
Tyr Glu Ile Leu Thr Pro Asn Thr Pro Phe Leu Ile Asp Asn Asp Gly
        185                 190                 195

aac att gag aac aca gag aag ctg cag tac agt ggt gag agg ctc tat   917
Asn Ile Glu Asn Thr Glu Lys Leu Gln Tyr Ser Gly Glu Arg Leu Tyr
    200                 205                 210

aag ttt aca gtg aca gct tat gac tgt ggg aag aag cgg gca gca gat   965
Lys Phe Thr Val Thr Ala Tyr Asp Cys Gly Lys Lys Arg Ala Ala Asp
215                 220                 225                 230

gat gct gag gtg gag att cag gtg aag ccc acc tgt aaa ccc agc tgg   1013
Asp Ala Glu Val Glu Ile Gln Val Lys Pro Thr Cys Lys Pro Ser Trp
                235                 240                 245

caa ggc tgg aac aaa agg atc gaa tat gca cca ggt gct ggg agc ttg   1061
Gln Gly Trp Asn Lys Arg Ile Glu Tyr Ala Pro Gly Ala Gly Ser Leu
            250                 255                 260

gct ttg ttc cct ggt atc cgc ctg gag acc tgt gat gaa cca ctc tgg   1109
Ala Leu Phe Pro Gly Ile Arg Leu Glu Thr Cys Asp Glu Pro Leu Trp
        265                 270                 275

aac att cag gcc acc ata gag ctg cag acc agc cat gtg gcc aag ggc   1157
Asn Ile Gln Ala Thr Ile Glu Leu Gln Thr Ser His Val Ala Lys Gly
    280                 285                 290

tgt gac cgt gac aac tac tca gag cgg gcg ctg cgg aaa ctc tgt ggt   1205
Cys Asp Arg Asp Asn Tyr Ser Glu Arg Ala Leu Arg Lys Leu Cys Gly
295                 300                 305                 310

gct gcc act ggg gag gtg gat ctg ttg ccc atg cct ggc ccc aat gcc   1253
Ala Ala Thr Gly Glu Val Asp Leu Leu Pro Met Pro Gly Pro Asn Ala
                315                 320                 325
```

```
aac tgg aca gca gga ctc tcg gtg cac tac agc cag gac agc agc ctg   1301
Asn Trp Thr Ala Gly Leu Ser Val His Tyr Ser Gln Asp Ser Ser Leu
            330                 335                 340

atc tac tgg ttc aat ggc acc cag gct gtg cag gtg ccc ctg ggt ggc   1349
Ile Tyr Trp Phe Asn Gly Thr Gln Ala Val Gln Val Pro Leu Gly Gly
        345                 350                 355

ccc agt ggg ctg ggc tct ggg ccc cag gac agc ctc agt gac cac ttc   1397
Pro Ser Gly Leu Gly Ser Gly Pro Gln Asp Ser Leu Ser Asp His Phe
    360                 365                 370

acc ctg tcc ttc tgg atg aag cat ggc gta act ccc aac aag ggc aag   1445
Thr Leu Ser Phe Trp Met Lys His Gly Val Thr Pro Asn Lys Gly Lys
375                 380                 385                 390

aag gaa gag gaa acc atc gta tgt aac act gtc cag aat gag gac ggc   1493
Lys Glu Glu Glu Thr Ile Val Cys Asn Thr Val Gln Asn Glu Asp Gly
                395                 400                 405

ttc tct cac tac tcg ctg act gtc cac ggc tgt agg att gcc ttc ctc   1541
Phe Ser His Tyr Ser Leu Thr Val His Gly Cys Arg Ile Ala Phe Leu
            410                 415                 420

tac tgg ccc ctg ctt gag agt gcc cgc cca gtc aag ttc ctc tgg aag   1589
Tyr Trp Pro Leu Leu Glu Ser Ala Arg Pro Val Lys Phe Leu Trp Lys
        425                 430                 435

ctg gag cag gtc tgt gat gat gag tgg cac cac tac gct ctg aac ctc   1637
Leu Glu Gln Val Cys Asp Asp Glu Trp His His Tyr Ala Leu Asn Leu
    440                 445                 450

gag ttc ccc aca gtc aca ctc tat acc gac ggc atc tcc ttc gac cct   1685
Glu Phe Pro Thr Val Thr Leu Tyr Thr Asp Gly Ile Ser Phe Asp Pro
455                 460                 465                 470

gcc ctc atc cat gac aat ggc ctc atc cac cca ccc cga agg gag cct   1733
Ala Leu Ile His Asp Asn Gly Leu Ile His Pro Pro Arg Arg Glu Pro
                475                 480                 485

gct ctc atg att ggg gcc tgc tgg act gag gag aag aac aaa gag aag   1781
Ala Leu Met Ile Gly Ala Cys Trp Thr Glu Glu Lys Asn Lys Glu Lys
            490                 495                 500

gaa aag gga gac aac agt aca gac acc acc caa gga gac cct ttg tcg   1829
Glu Lys Gly Asp Asn Ser Thr Asp Thr Thr Gln Gly Asp Pro Leu Ser
        505                 510                 515
```

```
atc cac cac tac ttc cat ggc tac ctg gct ggt ttc agc gtg cgc tca   1877
Ile His His Tyr Phe His Gly Tyr Leu Ala Gly Phe Ser Val Arg Ser
    520                 525                 530

ggt cgc ctg gag agc cgc gag gtc atc gag tgc ctc tat gca tgt cgg   1925
Gly Arg Leu Glu Ser Arg Glu Val Ile Glu Cys Leu Tyr Ala Cys Arg
535                 540                 545                 550

gag ggg ctg gac tat agg gat ttc gag agc ctg ggc aaa ggc atg aag   1973
Glu Gly Leu Asp Tyr Arg Asp Phe Glu Ser Leu Gly Lys Gly Met Lys
                555                 560                 565

gtc cac gtg aac ccc tca cag tcc ctg ctc acc ctg gag ggg gat gat   2021
Val His Val Asn Pro Ser Gln Ser Leu Leu Thr Leu Glu Gly Asp Asp
            570                 575                 580

gtg gag acc ttc aac cat gcc ctg cag cat gtg gct tac atg aac act   2069
Val Glu Thr Phe Asn His Ala Leu Gln His Val Ala Tyr Met Asn Thr
        585                 590                 595

ctg cgc ttt gcc acg ccc ggc gtc agg ccc ctg cgc ctc acc act gct   2117
Leu Arg Phe Ala Thr Pro Gly Val Arg Pro Leu Arg Leu Thr Thr Ala
    600                 605                 610

gtc aag tgc ttc agc gaa gag tcc tgc gtc tcc atc cct gaa gtg gag   2165
Val Lys Cys Phe Ser Glu Glu Ser Cys Val Ser Ile Pro Glu Val Glu
615                 620                 625                 630

ggc tac gtg gtc gtc ctt cag cct gac gcc ccc cag atc ctg ctg agt   2213
Gly Tyr Val Val Val Leu Gln Pro Asp Ala Pro Gln Ile Leu Leu Ser
                635                 640                 645

ggc act gct cat ttt gcc cgc cca gct gtg gac ttt gag gga acc aac   2261
Gly Thr Ala His Phe Ala Arg Pro Ala Val Asp Phe Glu Gly Thr Asn
            650                 655                 660

ggc gtc cct ttg ttc cct gat ctt caa atc acc tgc tcc att tct cac   2309
Gly Val Pro Leu Phe Pro Asp Leu Gln Ile Thr Cys Ser Ile Ser His
        665                 670                 675

cag gtg gag gcc aaa aag gat gag agt tgg cag ggc aca gtg aca gac   2357
Gln Val Glu Ala Lys Lys Asp Glu Ser Trp Gln Gly Thr Val Thr Asp
    680                 685                 690

aca cgc atg tcg gat gag att gtg cac aac ctg gat ggc tgt gaa att   2405
Thr Arg Met Ser Asp Glu Ile Val His Asn Leu Asp Gly Cys Glu Ile
695                 700                 705                 710
```

```
tct ctg gtg ggg gat gac ctg gat ccc gag cgg gaa agc ctg ctc ctg   2453
Ser Leu Val Gly Asp Asp Leu Asp Pro Glu Arg Glu Ser Leu Leu Leu
                715                 720                 725

gac aca acc tct ctg cag cag cgg ggg ctg gag ctc acc aac aca tct   2501
Asp Thr Thr Ser Leu Gln Gln Arg Gly Leu Glu Leu Thr Asn Thr Ser
            730                 735                 740

gcc tac ctc act att gct ggg gtg gag agc atc act gtg tat gaa gag   2549
Ala Tyr Leu Thr Ile Ala Gly Val Glu Ser Ile Thr Val Tyr Glu Glu
        745                 750                 755

atc ctg agg cag gct cgt tat cgg ctg cga cac gga gct gcc ctc tac   2597
Ile Leu Arg Gln Ala Arg Tyr Arg Leu Arg His Gly Ala Ala Leu Tyr
    760                 765                 770

acc agg aag ttc cgg ctt tcc tgc tcg gaa atg aat ggc cgt tac tcc   2645
Thr Arg Lys Phe Arg Leu Ser Cys Ser Glu Met Asn Gly Arg Tyr Ser
775                 780                 785                 790

agc aat gaa ttc atc gtg gag gtc aat gtc ctg cac agc atg aac cgg   2693
Ser Asn Glu Phe Ile Val Glu Val Asn Val Leu His Ser Met Asn Arg
                795                 800                 805

gtt gcc cac ccc agc cac gtg ctc agc tcc cag cag ttc ctg cac cgt   2741
Val Ala His Pro Ser His Val Leu Ser Ser Gln Gln Phe Leu His Arg
            810                 815                 820

ggt cac cag ccc ccg cct gag atg gct gga cac agc cta gcc agc tcc   2789
Gly His Gln Pro Pro Pro Glu Met Ala Gly His Ser Leu Ala Ser Ser
        825                 830                 835

cac aga aac tcc atg ata ccc agc gcc gca acc ctc atc att gtg gtg   2837
His Arg Asn Ser Met Ile Pro Ser Ala Ala Thr Leu Ile Ile Val Val
    840                 845                 850

tgc gtg ggc ttc ctg gtg ctc atg gtc gtc ctg ggc ctg gtg cgc atc   2885
Cys Val Gly Phe Leu Val Leu Met Val Val Leu Gly Leu Val Arg Ile
855                 860                 865                 870

cat tcc ctt cac cgc cgc gtc tca ggg gcc ggc ggg cct cca ggg gcc   2933
His Ser Leu His Arg Arg Val Ser Gly Ala Gly Gly Pro Pro Gly Ala
                875                 880                 885

tcc agt gac ccc aag gac cca gac ctc ttc tgg gat gac tca gct ctc   2981
Ser Ser Asp Pro Lys Asp Pro Asp Leu Phe Trp Asp Asp Ser Ala Leu
            890                 895                 900
```

```
acc atc att gtg aac ccc atg gag tcc tac cag aat cgg cag tcc tgt   3029
Thr Ile Ile Val Asn Pro Met Glu Ser Tyr Gln Asn Arg Gln Ser Cys
        905                 910                 915

gtg acg ggg gct gtt ggg ggc cag cag gag gat gag gac agc agt gac   3077
Val Thr Gly Ala Val Gly Gly Gln Gln Glu Asp Glu Asp Ser Ser Asp
    920                 925                 930

tcg gag gtg gcc gat tcc ccc agc agc gac gag aga cgc atc atc gag   3125
Ser Glu Val Ala Asp Ser Pro Ser Ser Asp Glu Arg Arg Ile Ile Glu
935                 940                 945                 950

acc ccc cca cac cgc tac taaggcctac acctctcccc acgcagaggg          3173
Thr Pro Pro His Arg Tyr
                955

ggaattctgc cctg                                                   3187
```

<210> 6
<211> 956
<212> PRT
<213> Homo sapiens

<400> 6

```
Met Thr Leu Leu Leu Leu Pro Leu Leu Leu Ala Ser Leu Leu Ala Ser
  1               5                  10                  15

Cys Ser Cys Asn Lys Ala Asn Lys His Lys Pro Trp Ile Glu Ala Glu
             20                  25                  30

Tyr Gln Gly Ile Val Met Glu Asn Asp Asn Thr Val Leu Leu Asn Pro
         35                  40                  45

Pro Leu Phe Ala Leu Asp Lys Asp Ala Pro Leu Arg Tyr Ala Gly Glu
     50                  55                  60

Ile Cys Gly Phe Arg Leu His Gly Ser Gly Val Pro Phe Glu Ala Val
 65                  70                  75                  80

Ile Leu Asp Lys Ala Thr Gly Glu Gly Leu Ile Arg Ala Lys Glu Pro
                 85                  90                  95

Val Asp Cys Glu Ala Gln Lys Glu His Thr Phe Thr Ile Gln Ala Tyr
            100                 105                 110

Asp Cys Gly Glu Gly Pro Asp Gly Ala Asn Thr Lys Lys Ser His Lys
        115                 120                 125
```

```
Ala Thr Val His Val Arg Val Asn Asp Val Asn Glu Phe Ala Pro Val
    130                 135                 140

Phe Val Glu Arg Leu Tyr Arg Ala Ala Val Thr Glu Gly Lys Leu Tyr
145                 150                 155                 160

Asp Arg Ile Leu Arg Val Glu Ala Ile Asp Gly Asp Cys Ser Pro Gln
                165                 170                 175

Tyr Ser Gln Ile Cys Tyr Tyr Glu Ile Leu Thr Pro Asn Thr Pro Phe
            180                 185                 190

Leu Ile Asp Asn Asp Gly Asn Ile Glu Asn Thr Glu Lys Leu Gln Tyr
        195                 200                 205

Ser Gly Glu Arg Leu Tyr Lys Phe Thr Val Thr Ala Tyr Asp Cys Gly
    210                 215                 220

Lys Lys Arg Ala Ala Asp Asp Ala Glu Val Glu Ile Gln Val Lys Pro
225                 230                 235                 240

Thr Cys Lys Pro Ser Trp Gln Gly Trp Asn Lys Arg Ile Glu Tyr Ala
                245                 250                 255

Pro Gly Ala Gly Ser Leu Ala Leu Phe Pro Gly Ile Arg Leu Glu Thr
            260                 265                 270

Cys Asp Glu Pro Leu Trp Asn Ile Gln Ala Thr Ile Glu Leu Gln Thr
        275                 280                 285

Ser His Val Ala Lys Gly Cys Asp Arg Asp Asn Tyr Ser Glu Arg Ala
    290                 295                 300

Leu Arg Lys Leu Cys Gly Ala Ala Thr Gly Glu Val Asp Leu Leu Pro
305                 310                 315                 320

Met Pro Gly Pro Asn Ala Asn Trp Thr Ala Gly Leu Ser Val His Tyr
                325                 330                 335

Ser Gln Asp Ser Ser Leu Ile Tyr Trp Phe Asn Gly Thr Gln Ala Val
            340                 345                 350

Gln Val Pro Leu Gly Gly Pro Ser Gly Leu Gly Ser Gly Pro Gln Asp
        355                 360                 365

Ser Leu Ser Asp His Phe Thr Leu Ser Phe Trp Met Lys His Gly Val
    370                 375                 380
```

```
Thr Pro Asn Lys Gly Lys Lys Glu Glu Glu Thr Ile Val Cys Asn Thr
385                 390                 395                 400

Val Gln Asn Glu Asp Gly Phe Ser His Tyr Ser Leu Thr Val His Gly
                405                 410                 415

Cys Arg Ile Ala Phe Leu Tyr Trp Pro Leu Leu Glu Ser Ala Arg Pro
            420                 425                 430

Val Lys Phe Leu Trp Lys Leu Glu Gln Val Cys Asp Asp Glu Trp His
        435                 440                 445

His Tyr Ala Leu Asn Leu Glu Phe Pro Thr Val Thr Leu Tyr Thr Asp
    450                 455                 460

Gly Ile Ser Phe Asp Pro Ala Leu Ile His Asp Asn Gly Leu Ile His
465                 470                 475                 480

Pro Pro Arg Arg Glu Pro Ala Leu Met Ile Gly Ala Cys Trp Thr Glu
                485                 490                 495

Glu Lys Asn Lys Glu Lys Glu Lys Gly Asp Asn Ser Thr Asp Thr Thr
            500                 505                 510

Gln Gly Asp Pro Leu Ser Ile His His Tyr Phe His Gly Tyr Leu Ala
        515                 520                 525

Gly Phe Ser Val Arg Ser Gly Arg Leu Glu Ser Arg Glu Val Ile Glu
    530                 535                 540

Cys Leu Tyr Ala Cys Arg Glu Gly Leu Asp Tyr Arg Asp Phe Glu Ser
545                 550                 555                 560

Leu Gly Lys Gly Met Lys Val His Val Asn Pro Ser Gln Ser Leu Leu
                565                 570                 575

Thr Leu Glu Gly Asp Asp Val Glu Thr Phe Asn His Ala Leu Gln His
            580                 585                 590

Val Ala Tyr Met Asn Thr Leu Arg Phe Ala Thr Pro Gly Val Arg Pro
        595                 600                 605

Leu Arg Leu Thr Thr Ala Val Lys Cys Phe Ser Glu Glu Ser Cys Val
    610                 615                 620

Ser Ile Pro Glu Val Glu Gly Tyr Val Val Val Leu Gln Pro Asp Ala
625                 630                 635                 640
```

Pro Gln Ile Leu Leu Ser Gly Thr Ala His Phe Ala Arg Pro Ala Val
645 650 655

Asp Phe Glu Gly Thr Asn Gly Val Pro Leu Phe Pro Asp Leu Gln Ile
660 665 670

Thr Cys Ser Ile Ser His Gln Val Glu Ala Lys Lys Asp Glu Ser Trp
675 680 685

Gln Gly Thr Val Thr Asp Thr Arg Met Ser Asp Glu Ile Val His Asn
690 695 700

Leu Asp Gly Cys Glu Ile Ser Leu Val Gly Asp Asp Leu Asp Pro Glu
705 710 715 720

Arg Glu Ser Leu Leu Leu Asp Thr Thr Ser Leu Gln Gln Arg Gly Leu
725 730 735

Glu Leu Thr Asn Thr Ser Ala Tyr Leu Thr Ile Ala Gly Val Glu Ser
740 745 750

Ile Thr Val Tyr Glu Glu Ile Leu Arg Gln Ala Arg Tyr Arg Leu Arg
755 760 765

His Gly Ala Ala Leu Tyr Thr Arg Lys Phe Arg Leu Ser Cys Ser Glu
770 775 780

Met Asn Gly Arg Tyr Ser Ser Asn Glu Phe Ile Val Glu Val Asn Val
785 790 795 800

Leu His Ser Met Asn Arg Val Ala His Pro Ser His Val Leu Ser Ser
805 810 815

Gln Gln Phe Leu His Arg Gly His Gln Pro Pro Pro Glu Met Ala Gly
820 825 830

His Ser Leu Ala Ser Ser His Arg Asn Ser Met Ile Pro Ser Ala Ala
835 840 845

Thr Leu Ile Ile Val Val Cys Val Gly Phe Leu Val Leu Met Val Val
850 855 860

Leu Gly Leu Val Arg Ile His Ser Leu His Arg Arg Val Ser Gly Ala
865 870 875 880

Gly Gly Pro Pro Gly Ala Ser Ser Asp Pro Lys Asp Pro Asp Leu Phe
885 890 895

Trp Asp Asp Ser Ala Leu Thr Ile Ile Val Asn Pro Met Glu Ser Tyr
900 905 910

Gln Asn Arg Gln Ser Cys Val Thr Gly Ala Val Gly Gly Gln Gln Glu
915 920 925

Asp Glu Asp Ser Ser Asp Ser Glu Val Ala Asp Ser Pro Ser Ser Asp
930 935 940

Glu Arg Arg Ile Ile Glu Thr Pro Pro His Arg Tyr
945 950 955

<210> 7
<211> 14
<212> PRT
<213> Gallus gallus

<400> 7
Ala Arg Val Asn Lys His Lys Pro Trp Ile Glu Thr Thr Tyr
1 5 10

<210> 8
<211> 24
<212> PRT
<213> Gallus gallus

<400> 8
His Lys Pro Trp Ile Glu Thr Thr Tyr His Gly Ile Val Thr Glu Asn
1 5 10 15

Asp Asn Thr Val Leu Leu Asp Pro
20

<210> 9
<211> 6
<212> PRT
<213> Gallus gallus

<400> 9
Val Glu Ala Val Asp Ala
1 5

<210> 10
<211> 17
<212> PRT
<213> Gallus gallus

<400> 10
Ile Glu Tyr Glu Pro Gly Thr Gly Ser Leu Ala Leu Phe Pro Ser Met
1 5 10 15

Arg

<210> 11
<211> 7
<212> PRT
<213> Gallus gallus

<400> 11
Ile Pro Asp Gly Val Val Thr
1 5

<210> 12
<211> 9
<212> PRT
<213> Gallus gallus

<400> 12
Thr Tyr Lys Pro Ala Glu Phe His Trp
1 5

<210> 13
<211> 11
<212> PRT
<213> Gallus gallus

<400> 13
Glu Gly Leu Asp Leu Gln Ile Ala Asp Gly Val
1 5 10

<210> 14
<211> 28
<212> PRT
<213> Gallus gallus

<400> 14
Gly Ile Glu Met Ser Ser Ser Asn Leu Gly Met Ile Ile Thr Gly Val
1 5 10 15

Asp Thr Met Ala Ser Tyr Glu Glu Val Leu His Leu
20 25

<210> 15
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 15
gtaaamaagc ayaagccatg gat 23

<210> 16
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 16
caggaathgt aacagagaat gataa 25

<210> 17
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 17
ccagtaccag gctcatactc dat 23

<210> 18
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 18
gtatcaacac catadatdat catacc 26

<210> 19
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 19
acaccatcag cdatctgaaa atc 23

<210> 20
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 20
gcatcaaact cagcctcctt ataaaa 26

<210> 21
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 21
ggggaacaga agagctgcac atcagcgaac g 31

<210> 22
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 22
ccccctcgag ttagtagctg agtgtggag 29

<210> 23
<211> 39
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 23
gggccatggc tcgtgttaac aagcataagc cctggattg 39

<210> 24
<211> 47
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 24
cccaagctta gtggtggtgg tgatggtgtg gttcatcaca tgtgtcc 47

<210> 25
<211> 268
<212> PRT
<213> Gallus gallus

<400> 25
Ala Arg Val Asn Lys His Lys Pro Trp Ile Glu Thr Thr Tyr His Gly
1 5 10 15

Ile Val Thr Glu Asn Asp Asn Thr Val Leu Leu Asp Pro Pro Leu Ile
20 25 30

Ala Leu Asp Lys Asp Ala Pro Leu Arg Phe Ala Glu Ser Phe Glu Val
35 40 45

Thr Val Thr Lys Glu Gly Glu Ile Cys Gly Phe Leu Lys Ile His Gly

50 55 60

Gln Asn Val Pro Phe Glu Ala Val Val Val Asp Lys Ser Thr Gly Glu
65 70 75 80

Gly Ile Ile Arg Ser Lys Glu Lys Leu Asp Cys Glu Leu Gln Lys Asp
85 90 95

Tyr Thr Phe Thr Ile Gln Ala Tyr Asp Cys Gly Lys Gly Pro Asp Gly
100 105 110

Ala Asn Ala Lys Lys Ser His Lys Ala Thr Val His Ile Gln Val Asn
115 120 125

Asp Val Asn Glu Tyr Ser Pro Val Phe Lys Glu Lys Ser Tyr Lys Ala
130 135 140

Thr Val Ile Glu Gly Lys Arg Tyr Asp Asn Ile Leu Lys Val Glu Ala
145 150 155 160

Val Asp Ala Asp Cys Ser Pro Gln Phe Ser Gln Ile Cys Asn Tyr Glu
165 170 175

Ile Val Thr Pro Asp Val Pro Phe Ala Ile Asp Lys Asp Gly Tyr Ile
180 185 190

Lys Asn Thr Glu Lys Leu Ser Tyr Gly Lys Glu His Gln Tyr Lys Leu
195 200 205

Thr Val Thr Ala Tyr Asp Cys Gly Lys Lys Arg Ala Ala Glu Asp Val
210 215 220

Leu Val Lys Ile Ser Ile Lys Pro Thr Cys Lys Pro Gly Trp Gln Gly
225 230 235 240

Trp Ser Lys Arg Ile Glu Tyr Glu Pro Gly Thr Gly Ser Leu Ala Leu
245 250 255

Phe Pro Ser Met Arg Leu Glu Thr Cys Asp Glu Pro
260 265

<210> 26
<211> 29
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:primer

<400> 26
gggccatgat acgctacaga aactggcac 29

<210> 27
<211> 49
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:primer

<400> 27
cccaagctta gtggtggtgg tgatggtgag tggctgtact tggaacaac 49

<210> 28
<211> 87
<212> PRT
<213> Gallus gallus

<400> 28
Ile Arg Tyr Arg Asn Trp His Thr Val Ser Leu Phe Asp Arg Lys Phe
1 5 10 15

Lys Leu Val Cys Ser Glu Leu Asn Gly Arg Tyr Val Ser Asn Glu Phe
20 25 30

Lys Val Glu Val Asn Val Ile His Thr Ala Asn Pro Ile Glu His Ala
35 40 45

Asn His Ile Ala Ala Gln Pro Gln Phe Val His Pro Val His His Thr
50 55 60

Phe Val Asp Leu Ser Gly His Asn Leu Ala Asn Pro His Pro Phe Ser
65 70 75 80

Val Val Pro Ser Thr Ala Thr
85

<210> 29
<211> 89
<212> PRT
<213> Gallus gallus

<400> 29
Leu Ile Arg Tyr Arg Asn Trp His Thr Val Ser Leu Phe Asp Arg Lys
1 5 10 15

Phe Lys Leu Val Cys Ser Glu Leu Asn Gly Arg Tyr Val Ser Asn Glu
20 25 30

Phe Lys Val Glu Val Asn Val Ile His Thr Ala Asn Pro Ile Glu His
35 40 45

Ala Asn His Ile Ala Ala Gln Pro Gln Phe Val His Pro Val His His
50 55 60

Thr Phe Val Asp Leu Ser Gly His Asn Leu Ala Asn Pro His Pro Phe
65 70 75 80

Ser Val Val Pro Ser Thr Ala Thr Val
85

<210> 30
<211> 89
<212> PRT
<213> Gallus gallus

<400> 30
Leu Leu Arg Tyr Arg Asn Trp His Ala Arg Ser Leu Leu Asp Arg Lys
1 5 10 15

Phe Lys Leu Ile Cys Ser Glu Leu Asn Gly Arg Tyr Ile Ser Asn Glu
20 25 30

Phe Lys Val Glu Val Asn Val Ile His Thr Ala Asn Pro Met Glu His
35 40 45

Ala Asn His Met Ala Ala Gln Pro Gln Phe Val His Pro Glu His Arg
50 55 60

Ser Phe Val Asp Leu Ser Gly His Asn Leu Ala Asn Pro His Pro Phe
65 70 75 80

Ala Val Val Pro Ser Thr Ala Thr Val
85

<210> 31
<211> 22
<212> DNA

<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 31
ctcctctggc atcattgacc tc 22

<210> 32
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 32
catttcttcc tcggcttctt cc 22

<210> 33
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 33
tgctgcgagg atgctgc 17

<210> 34
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 34
atgatgccag aggaggc 17

<210> 35
<211> 85
<212> PRT

<213> Homo sapiens

<400> 35

His	Ile	Arg	Tyr	Arg	Asn	Trp	Arg	Pro	Ala	Ser	Leu	Glu	Ala	Arg	Arg
1				5				10						15	

Phe	Arg	Ile	Lys	Cys	Ser	Glu	Leu	Asn	Gly	Arg	Tyr	Thr	Ser	Asn	Glu
		20					25					30			

Phe	Asn	Leu	Glu	Val	Ser	Ile	Leu	His	Glu	Asp	Gln	Val	Ser	Asp	Lys
	35					40					45				

Glu	His	Val	Asn	His	Leu	Ile	Val	Gln	Pro	Pro	Phe	Leu	Gln	Ser	Val
	50				55					60					

His	His	Pro	Glu	Ser	Arg	Ser	Ser	Ile	Gln	His	Ser	Ser	Val	Val	Pro
65				70					75						80

Ser	Ile	Ala	Thr	Val
			85	

<210> 36
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 36
ctgcagtagc ggggttg 17

<210> 37
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 37
tggagtgtct gtttcaccag g 21

<210> 38
<211> 87

<212> PRT
<213> Homo sapiens

<400> 38
Ile Leu Arg Gln Ala Arg Tyr Arg Leu Arg His Gly Ala Ala Leu Tyr
1 5 10 15

Thr Arg Lys Phe Arg Leu Ser Cys Ser Glu Met Asn Gly Arg Tyr Ser
20 25 30

Ser Asn Glu Phe Ile Val Glu Val Asn Val Leu His Ser Met Asn Arg
35 40 45

Val Ala His Pro Ser His Val Leu Ser Ser Gln Gln Phe Leu His Arg
50 55 60

Gly His Gln Pro Pro Pro Glu Met Ala Gly His Ser Leu Ala Ser Ser
65 70 75 80

His Arg Asn Ser Met Ile Pro
85

<210> 39
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PCR Primer

<400> 39
cgggatcccg catccgggcc gcacat 26

<210> 40
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PCR Primer

<400> 40
gggaattcct cagtagctga gggtggag 28

<210> 41

<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PCR Primer

<400> 41
ccggaattcc cttctgatgg ggaccatcag 30

<210> 42
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PCR Primer

<400> 42
ccgctcgagt cagtcttccc actcatcatc 30